

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 20:11:57 ; Search time 7402 Seconds
(without alignments)
11712.680 Million cell updates/sec

Title: US-09-842-484A-1
Perfect score: 2979
Sequence: 1 ttataactgattaaagaag.....acatttgattttattaaaa 2979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_vl.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_fun.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2979	100.0	2979	1	AF195517	AF195517 Pasteurel
2	2635.8	88.5	11885	1	AE006116	AE006116 Pasteurel
3	2632.6	88.4	8838	1	AF302467	AF302467 Pasteurel
4	2340.8	78.6	16727	1	AF067175	AF067175 Pasteurel
5	2283.4	76.6	2937	1	AF036004	AF036004 Pasteurel
6	2268.6	76.2	2919	1	AF237926	AF237926 Pasteurel
c 7	814.2	27.3	14483	1	AB079602	AB079602 Escherich
c 8	97.6	3.3	155204	2	AC007926	AC007926 Trypanoso
c 9	86.2	2.9	100726	2	AC116961	AC116961 Dictyoste
c 10	82.6	2.8	43993	2	AC116965	AC116965 Dictyoste
c 11	82.4	2.8	11442	1	AF400048	AF400048 Campyloba
c 12	82.4	2.8	11455	1	AY044868	AY044868 Campyloba
c 13	82.4	2.8	16265	1	VCLPSS	Y07786 V.cholerae
14	82.4	2.8	46721	1	AB012956	AB012956 Vibrio ch
15	82.4	2.8	270050	1	AL591977	AL591977 Listeria
c 16	81.8	2.7	153026	9	AL391280	AL391280 Human DNA
c 17	81.6	2.7	54345	3	AC084152	AC084152 Caenorhab
c 18	81.6	2.7	302000	1	AP003187	AP003187 Clostridi
19	81	2.7	984	6	AX413863	AX413863 Sequence
20	81	2.7	984	6	AX415777	AX415777 Sequence
c 21	80.8	2.7	12388	1	AF401529	AF401529 Campyloba
c 22	80.8	2.7	12390	1	AF401528	AF401528 Campyloba
c 23	80.8	2.7	15671	1	AF343914	AF343914 Campyloba
24	80.6	2.7	9987	1	AB050723	AB050723 Streptoco
25	80.6	2.7	62352	2	AC116990	AC116990 Dictyoste
c 26	80.6	2.7	171798	2	AC084397	AC084397 Trypanoso
27	80.4	2.7	18239	1	AF349539	AF349539 Streptoco
c 28	80.4	2.7	86827	3	PFMAL3P5	AL034556 Plasmodiu
c 29	80.4	2.7	347400	1	AP003591	AP003591 Nostoc sp
c 30	80	2.7	19966	6	AF373595	AF373595 Streptoco
c 31	80	2.7	19966	6	AX009404	AX009404 Sequence
32	80	2.7	27782	1	AF448502	AF448502 Streptoco
c 33	80	2.7	178273	2	AC005308	AC005308 Plasmodiu
c 34	79.4	2.7	10827	1	AE006155	AE006155 Pasteurel
c 35	79.2	2.7	840	8	CNS0180K	AL110675 Botrytis
c 36	79.2	2.7	11474	1	AF130984	AF130984 Campyloba
c 37	79.2	2.7	11474	1	AF167344	AF167344 Campyloba
c 38	79.2	2.7	11474	1	AF215659	AF215659 Campyloba
c 39	79	2.7	20416	3	AC114257	AC114257 Dictyoste
c 40	79	2.7	107289	2	AC116923	AC116923 Dictyoste
c 41	79	2.7	158548	3	PFMAL3P2	AL034558 Plasmodiu
c 42	78.6	2.6	106434	3	AC117080	AC117080 Dictyoste
c 43	78.2	2.6	12728	2	AC116101	AC116101 Dictyoste
c 44	78	2.6	16563	2	AC115600	AC115600 Dictyoste
45	78	2.6	104992	2	AC005504	AC005504 Plasmodiu

ALIGNMENTS

RESULT 1
AF195517
LOCUS AF195517 Pasteurella multocida chondroitin synthase CS gene, complete cds.
DEFINITION Pasteurella multocida.
ACCESSION AF195517
VERSION AF195517.1 GI:9716369
KEYWORDS
SOURCE Pasteurella multocida.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
REFERENCE 1 (bases 1 to 2979)
AUTHORS DeAngelis,P.L. and Padgett-McCue,A.J.
TITLE Identification and molecular cloning of a chondroitin synthase from Pasteurella multocida type F

J. Biol. Chem. 275 (31), 24124-24129 (2000)
20379058
10818104
PUBMED
2 (bases 1 to 2979)
DeAngelis, P.L. and Padgett-McCue, A.J.
Direct Submission
Submitted (15-OCT-1999) Biochem. & Molec. Biol., Univ. of Oklahoma
Health Sciences Center, 940 Stanton L. Young Blvd., Oklahoma City,
OK 73104, USA
FEATURES
source
1..2979
/organism="Pasteurella multocida"
/strain="p4679"
/db_xref="taxon:747"
/country="USA: Arkansas"
/note="capsule type F"
isolated from face lesion of turkey from Arkansas; strain
obtained from USDA
61..2958
/function="synthesizes the chondroitin polymer backbone (a
polysaccharide composed of repeating disaccharide units
consisting of beta(1,4)glucuronic acid linked to
beta(1,3)N-acetylgalactosamine) involved in capsule
biosynthesis"
/note="polysaccharide-forming glycosyltransferase; similar
to Pasteurella multocida type A hyaluronan synthase"
/codon_start=1
/transl_table=11
/product="chondroitin synthase CS"
/protein_id="AAF97500.1"
/db_xref="GI:9716370"
/translation="MPTLSQAIKAYNSDYELAKLPEKSAETVGRKIVFQIKCKE
KLSTNSVSEDKNSVCDSDIATQLLLSNVKKLISESEKSLNKKWSIKNGKSE
NATRIRKVELPKDFPDLVLPDHWNDFTWKNRKKSGIKRPVKNIGLSIILP
NBSRIIDITLACLVNQKNTYFFVADVDDSKENLLTVOKYEQKLDIKVROKNDGY
QICAVRNLGRRTAKYDFVSLDQMDAPQOLWHSYLETLEDNDIVLIGPRKVDTHN
ITAEQFLNDPYLESTPETATNNPSITSGNISLWRLHFRKTDNRLCDSPPFRF
VAGNFAEKLKNGKWFDEENHGWGDEVEFGVRLFAKGFVRVIDGMAIHOEPPG
KENETERAKSKTLKIVKVPVYRKLLPIEDSHTRIPLVSIYIPAYNCANYIOR
CVSALNQTLDVLCICNGSDTNTLEVLNKLXGNNPRVIRMSKPNKGSIASNAVA
SFAGKYIIGDSDYLEPDAVELCLKEFLDKLILACVYTINRVNPDGSLIANGYN
PEFSREKLTITAMTAHFMETIRAWHLTDGFNEIENAVDIDMLKUSEVGKFKHLNK
ICNVRLHGDNDSTKLGIQKNHVVVNSLNROGINYNKDFDLDESRRYIFNK
TAYOEEMDMKDLKJLONKDAKIAVSIFYPNTLNGVLKLNINIEYNKNIPLVILHV
DKNHLTPDIKKEILAFYKHQVNNILLNNDISYTSNRLIKTEAHLNINLSQNLNC
EYIFDNHDSLFVYKNDYVYMKYVDGNFSAETHDLEKINAHPPPKKLIKTYFNDN
DLRSMNVKASQGMFMKIALPHELITIKIVETSCQSIDSVPEYNTEDINQFALLIL
EKKTGHVFNKTSITLTYMPWERKLOWTNEQIOSARKGENIPVKNFIINSITL"
BASE COUNT 1129 a 466 c 497 g 887 t
ORIGIN
Query Match 100.08; Score 2979; DB 1; Length 2979;
Best Local Similarity 100.08; Pred. No. 0;
Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAAAGTAAAGAGGTAAGGATTAAGCAAGGTTAATTTTAAAGGAAGAAA 60
DB 1 TTATAAAGTAAAGAGGTAAGGATTAAGCAAGGTTAATTTTAAAGGAAGAAA 60
QY 61 ATGAATACATATACCAAGCAATAAAGCATATACAGCAAGTACATGATGACATC 120
DB 61 ATGAATACATATACCAAGCAATAAAGCATATACAGCAAGTACATGATGACATC 120
QY 121 AAATTTATTTGAGAAGTCTGCTGAAAGCTACGGCGGAGAAAATCGTTGAATTC 180
DB 121 AAATTTATTTGAGAAGTCTGCTGAAAGCTACGGCGGAGAAAATCGTTGAATTC 180
QY 181 AAATGTAAGAAAAAAGTCTCGACCAATTCCTATGTAAGTGAAGATTAAGAAACAGTGT 240
DB 181 AAATGTAAGAAAAAAGTCTCGACCAATTCCTATGTAAGTGAAGATTAAGAAACAGTGT 240
QY 241 TGGATAGCTCATAGATTCGCAACACAGCTCTTACTTTCCCAACGTAAGAAAAATTAAC 300

241 TCGATAGCTCATAGATTCGCAACACAGCTCTTACTTTCCCAACGTAAGAAAAATTAAC 300
QY 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTACTCTGGGAAAAAA 360
DB 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTACTCTGGGAAAAAA 360
QY 361 TCGGAGAACGCAAGAAATCAGAAAGGTGGAACTAGTACCCAAAGATTTTCTAAGATCTT 420
DB 361 TCGGAGAACGCAAGAAATCAGAAAGGTGGAACTAGTACCCAAAGATTTTCTAAGATCTT 420
QY 421 GTTCTTGCTCCATGGCAGATCATGTTAATGATTTACATGTTACATGTTACAAAAATCGAAAA 480
DB 421 GTTCTTGCTCCATGGCAGATCATGTTAATGATTTACATGTTACATGTTACAAAAATCGAAAA 480
QY 481 AGCTTAGGTATAAGCCTGTAAATAAGATATCGGTCTTCTTATTTATTTTCTTACATTT 540
DB 481 AGCTTAGGTATAAGCCTGTAAATAAGATATCGGTCTTCTTATTTATTTTCTTACATTT 540
QY 541 AATCGTAGCCGTATTTTAGATATAACGTTAGCCGTGTTGGTCAATCAGAAAAACAACTAC 600
DB 541 AATCGTAGCCGTATTTTAGATATAACGTTAGCCGTGTTGGTCAATCAGAAAAACAACTAC 600
QY 601 CCATTTGAAGTCGTTGTCAGATGATGGTAGTAAAGAAAACTTACTTACCATTGTGCAA 660
DB 601 CCATTTGAAGTCGTTGTCAGATGATGGTAGTAAAGAAAACTTACTTACCATTGTGCAA 660
QY 661 AAATACGAACAAAAAATTTGACATAAAGTATGTAAGACAAAAAAGATTTATGGATATCAATTG 720
DB 661 AAATACGAACAAAAAATTTGACATAAAGTATGTAAGACAAAAAAGATTTATGGATATCAATTG 720
QY 721 TGTGCAAGTACAGAACTTAGGTTTACGTACACAAAGTATGTTGTCGATTTCTAGAC 780
DB 721 TGTGCAAGTACAGAACTTAGGTTTACGTACACAAAGTATGTTGTCGATTTCTAGAC 780
QY 781 TGGATATGGCACCACAACTATGGTTCATTCTTATCTTACAGAACTATTAGAAGAC 840
DB 781 TGGATATGGCACCACAACTATGGTTCATTCTTATCTTACAGAACTATTAGAAGAC 840
QY 841 AATGATATGTTTAAATTTGGACCTAGAAAAATATGCGTACTCATTAATATTACCCAGAA 900
DB 841 AATGATATGTTTAAATTTGGACCTAGAAAAATATGCGTACTCATTAATATTACCCAGAA 900
QY 901 CAATTCCTTAAGATCCCATATTAAGATACACCTACCTGAAACCGCTACAATTAACAT 960
DB 901 CAATTCCTTAAGATCCCATATTAAGATACACCTACCTGAAACCGCTACAATTAACAT 960
QY 961 CTTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAA 1020
DB 961 CTTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAA 1020
QY 1021 ACCGATATCTACGCTATGCTATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTGCA 1080
DB 1021 ACCGATATCTACGCTATGCTATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTGCA 1080
QY 1081 TTTTCTTAAAGATGGCTAAATAAAGTAGTGTGGTTCGATGAAGAAATTAATCATTTGGGG 1140
DB 1081 TTTTCTTAAAGATGGCTAAATAAAGTAGTGTGGTTCGATGAAGAAATTAATCATTTGGGG 1140
QY 1141 GCGGAAGATGTAAGTTGGTTACAGATTATTTGCCAAAGGCTGTTTTTTCAGAGTAATT 1200
DB 1141 GCGGAAGATGTAAGTTGGTTACAGATTATTTGCCAAAGGCTGTTTTTTCAGAGTAATT 1200
QY 1201 GACGGCGGAATGCCATCCATCAAGAACCCACCTGGTAAAGAAAAATGAACAGAACGGA 1260
DB 1201 GACGGCGGAATGCCATCCATCAAGAACCCACCTGGTAAAGAAAAATGAACAGAACGGA 1260
QY 1261 GCTGGTAAAGATTAACGCTTAAAAATTTGTGAAAGAAAGTACCTTACATCTATAGAAG 1320
DB 1261 GCTGGTAAAGATTAACGCTTAAAAATTTGTGAAAGAAAGTACCTTACATCTATAGAAG 1320
QY 1321 CTTTACCATAGAGATTCACATATTCATAGAAATACCTTTAGTTTCTTATTTATATCCCC 1380
DB 1321 CTTTACCATAGAGATTCACATATTCATAGAAATACCTTTAGTTTCTTATTTATATCCCC 1380

```
QY 1381 GCTTAAACCTGTGCAAAATTATATTCAAAGATGCTAGATAGTCTCTTAATCAAACCTGTT 1440
|||||
Db 1381 GCTTAAACCTGTGCAAAATTATATTCAAAGATGCTAGATAGTCTCTTAATCAAACCTGTT 1440
|||||
QY 1441 GTCGATCTCGAGGTTGTGATTTGTAAGATGGTTCAACAGATATAATACCTTTAGAGTGATC 1500
|||||
Db 1441 GTCGATCTCGAGGTTGTGATTTGTAAGATGGTTCAACAGATATAATACCTTTAGAGTGATC 1500
|||||
QY 1501 AATAAGCTTTATGGTATATATCTAGGTTACGATCATGTCTAAACCAATGGCGGAATA 1560
|||||
Db 1501 AATAAGCTTTATGGTATATATCTAGGTTACGATCATGTCTAAACCAATGGCGGAATA 1560
|||||
QY 1561 GCCTCAGCATCAAAATGCAGCGCTTCTCTTTGCTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
|||||
Db 1561 GCCTCAGCATCAAAATGCAGCGCTTCTCTTTGCTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
|||||
QY 1621 TCAGATGATTATCTTGAGCGTGTAGTCAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT 1680
|||||
Db 1621 TCAGATGATTATCTTGAGCGTGTAGTCAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT 1680
|||||
QY 1681 AAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTCAATCCGGATGGTGTATATC 1740
|||||
Db 1681 AAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTCAATCCGGATGGTGTATATC 1740
|||||
QY 1741 GCTAATGGTTACAATTTGCCAGAAATTTTACAGAGAAAACTCAACACGGCTATGATTGCT 1800
|||||
Db 1741 GCTAATGGTTACAATTTGCCAGAAATTTTACAGAGAAAACTCAACACGGCTATGATTGCT 1800
|||||
QY 1801 CACCATTTTGAATGTTTACGATPAGAGCTTGGCATTTAACCGATGGATTTAACGAAAT 1860
|||||
Db 1801 CACCATTTTGAATGTTTACGATPAGAGCTTGGCATTTAACCGATGGATTTAACGAAAT 1860
|||||
QY 1861 ATTGAACCGCGTGGATTTATGACATCTTCCCTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
|||||
Db 1861 ATTGAACCGCGTGGATTTATGACATCTTCCCTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
|||||
QY 1921 CATCTTAAATAAATCTGCTATAACCGCGTATTACATGGTGATACACATCCATTAAGAAA 1980
|||||
Db 1921 CATCTTAAATAAATCTGCTATAACCGCGTATTACATGGTGATACACATCCATTAAGAAA 1980
|||||
QY 1981 CTGGCATTCAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
|||||
Db 1981 CTGGCATTCAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
|||||
QY 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGATGAACTAGAAAGTATATCTTC 2100
|||||
Db 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGATGAACTAGAAAGTATATCTTC 2100
|||||
QY 2101 AATAAACCCTGTAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
|||||
Db 2101 AATAAACCCTGTAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
|||||
QY 2161 AATAAGATGCCAAATTCGAGTCAGTATTTTCTATCCCAATACATATAACGGCTTAGTG 2220
|||||
Db 2161 AATAAGATGCCAAATTCGAGTCAGTATTTTCTATCCCAATACATATAACGGCTTAGTG 2220
|||||
QY 2221 AAAAACTAACAATATTATTGATATAATAAATAATATATCGTTATATCTCATGTT 2280
|||||
Db 2221 AAAAACTAACAATATTATTGATATAATAAATAATATATCGTTATATCTCATGTT 2280
|||||
QY 2281 GATAAGAAATCATCTTACACAGACATCAAAAAAGAAATATTTGGCTTTCTATCATAAAGCAC 2340
|||||
Db 2281 GATAAGAAATCATCTTACACAGACATCAAAAAAGAAATATTTGGCTTTCTATCATAAAGCAC 2340
|||||
QY 2341 CAAGTGAATATTTTACTAAATAATGACATCTCATATTTACAGAGTAATAGACTAATAAAA 2400
|||||
Db 2341 CAAGTGAATATTTTACTAAATAATGACATCTCATATTTACAGAGTAATAGACTAATAAAA 2400
|||||
QY 2401 ACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
|||||
Db 2401 ACTGAGGCACATTTAAGTAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTGAATAC 2460
|||||
```

```
QY 2461 ATCATTTTTGATAATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAA 2520
|||||
Db 2461 ATCATTTTTGATAATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAA 2520
|||||
QY 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAAT 2580
|||||
Db 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAAT 2580
|||||
QY 2581 GCGCATCCACCATTTAAAAAGCTGATTAACACCTTATTTTAAATGACAAATGACTTAAGAAGT 2640
|||||
Db 2581 GCGCATCCACCATTTAAAAAGCTGATTAACACCTTATTTTAAATGACAAATGACTTAAGAAGT 2640
|||||
QY 2641 ATGAATGTGAAAGGGGCATCAACAGGTATGTTTATGAAGTATGGCTACCCGCATGAGCTT 2700
|||||
Db 2641 ATGAATGTGAAAGGGGCATCAACAGGTATGTTTATGAAGTATGGCTACCCGCATGAGCTT 2700
|||||
QY 2701 CTGACGATTTAATAAGAGTCATCACATCTCGGCCAAATCAATGATAGTGTGCCAGAATAT 2760
|||||
Db 2701 CTGACGATTTAATAAGAGTCATCACATCTCGGCCAAATCAATGATAGTGTGCCAGAATAT 2760
|||||
QY 2761 AACACTGAGGATATTTGGTTCCAAATTTGCCACTTTTAACTTTAGAAAAAGAAACCGCCAT 2820
|||||
Db 2761 AACACTGAGGATATTTGGTTCCAAATTTGCCACTTTTAACTTTAGAAAAAGAAACCGCCAT 2820
|||||
QY 2821 GTATTTAATAAACATCGACCTGACTTATATATGCTTTGGGAACGAAAAATTACAATGGACA 2880
|||||
Db 2821 GTATTTAATAAACATCGACCTGACTTATATATGCTTTGGGAACGAAAAATTACAATGGACA 2880
|||||
QY 2881 AATGAACAAATTCAAAGTGCAAAAAGGCGGAAAAATATCCCGTTTAAACAAGTTCAATTT 2940
|||||
Db 2881 AATGAACAAATTCAAAGTGCAAAAAGGCGGAAAAATATCCCGTTTAAACAAGTTCAATTT 2940
|||||
QY 2941 AATAGTATAACGCTATAAACACATTTGCAATTTTATTAATAA 2979
|||||
Db 2941 AATAGTATAACGCTATAAACACATTTGCAATTTTATTAATAA 2979
|||||

RESULT 2
AE006116 11885 bp DNA linear BCT 08-MAR-2001
LOCUS Pasteurella multocida PM70 section 83 of 204 of the complete
DEFINITION genome.
ACCESSION AE006116 AE004439
VERSION AE006116.1 GI:12721075
KEYWORDS
SOURCE Pasteurella multocida.
ORGANISM Pasteurella multocida.
REFERENCE 1 (bases 1 to 11885)
AUTHORS May, B. J., Zhang, Q., Li, L. L., Paustian, M. L., Whittam, T. S. and
Kapur, V.
TITLE Complete genomic sequence of Pasteurella multocida, Pm70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE 21145866
PUBMED 11248100
REFERENCE 2 (bases 1 to 11885)
AUTHORS Zhang, Q. and Kapur, V.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA

FEATURES
source 1..11885
/organism="Pasteurella multocida"
/strain="PM70"
/db_xref="taxon:747"
205..2073
/gene="hyaE"
/notes="PM0774"
205..2073
/gene="hyaE"
/codon_start=1

gene
CDS
```


/product="unknown"
/protein_id="AAK02866.1"
/db_xref="GI:12721084"
/translation="METLDKIKKQISENPILLYMKSPKFPSCGFSARAVEALMHCKV
PGFYVDILQHPDIRAELPAYANWTFPQLWVGDELVGGCDIILEMFQQGLQLLLADV
AKYQGE"
BASE COUNT 4287 a 2084 c 1962 g 3552 t
ORIGIN

Query Match 88.5%; Score 2635.8; DB 1; Length 11885;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 271; Conservative 0; Mismatches 207; Indels 1; Gaps 1;

QY	1	TTATAACTGATTAAAGAAGGTAACAGATTCAAGCRAAGGTTAAATTTTAAAGGAAAGAAA	60
DB	2031	TTATAAACTGATTAAAGAAGGTAACAGATTCAAGCRAAGGTTAAATTTT - AAGGAAGAAA	2089
QY	61	ATCAATACATTTATCAGAACGAAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC	120
DB	2090	ATGAATACATTTATCAGAACGAAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC	2149
QY	121	AAATTTATTCAGAAAGTCTGCTGAACCTACGGCGGAAATAATCGTTGAATTCCTCAAAATTATC	180
DB	2150	AAATTTATTCAGAAAGTCTGCTGAACCTACGGCGGAAATAATCGTTGAATTTCCAAATTTATC	2209
QY	181	AAATGTAAGAAAGAACTCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGTT	240
DB	2210	AAATGTAAGAAAGAACTCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGTT	2269
QY	241	TGGCATAGCTCATTAGATATCGCAACACAGCTCTTACTTTCCACGTAAGAAATTAAC	300
DB	2270	TGGCATAGCTCATTAGATATCGCAACACAGCTCTTACTTTCCACGTAAGAAATTAAC	2329
QY	301	CTATCCGAATCAGAAAAACAGTTTAAAAATAAAGTGAAGTATCATCTACCTGGGAAAAA	360
DB	2330	CTATCCGAATCAGAAAAACAGTTTAAAAATAAAGTGAAGTATCATCTACCTGGGAAAAA	2389
QY	361	TCGGAGAACGCAAGAAATCAGAAAGGTGGAAGTACTAGTACCCCAAGATTTTCCATAAGATCTT	420
DB	2390	TCGGAGAACGCAAGAAATCAGAAAGGTGGAAGTACTAGTACCCCAAGATTTTCCATAAGATCTT	2449
QY	421	GTCTTCTGCTCCATTGCCAGATCATGTTAAATGATTTTACATGGTACAAAAATCGAAAAA	480
DB	2450	GTCTTCTGCTCCATTGCCAGATCATGTTAAATGATTTTACATGGTACAAAAATCGAAAAA	2509
QY	481	AGCTTAGGTATAAGCCTGTAATAAGAAATATCGTCTTTCTTATTATTCTCTACATTT	540
DB	2510	AGCTTAGGTATAAGCCTGTAATAAGAAATATCGTCTTTCTTATTATTCTCTACATTT	2569
QY	541	AATCGTAGCCGATTTTATAGATATAACGTTAGCCCTTGGTTCATCAGAAAAACAACTAC	600
DB	2570	AATCGTAGCCGATTTTATAGATATAACGTTAGCCCTTGGTTCATCAGAAAAACAACTAC	2629
QY	601	CCATTTGAAGTCGTTGTGCAGATGATGGTAGTAAGAAAACTTACATTACATTGTGCA	660
DB	2630	CCATTTGAAGTCGTTGTGCAGATGATGGTAGTAAGAAAACTTACATTACATTGTGCA	2689
QY	661	AAATACGAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGATATCAATTG	720
DB	2690	AAATACGAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGATATCAATTG	2749
QY	721	TGTGCAGTCAGAACTTATAGTTTACGTACAGCAAGTATGATTTTGTCTCGATTTCTAGAC	780
DB	2750	TGTGCAGTCAGAACTTATAGTTTACGTACAGCAAGTATGATTTTGTCTCGATTTCTAGAC	2809
QY	781	TGGCATATGGCACACAACTATATGGTTTCAATTTCTTATCTTACAGAACTATTAGAGAC	840
DB	2810	TGGCATATGGCACACAACTATATGGTTTCAATTTCTTATCTTACAGAACTATTAGAGAC	2859
QY	841	AATGATATTTCTTTAATTGGACCTAGAAAAATATGTGGATACTCATATATTACCGCAGAA	900
DB	2870	AATGATATTTCTTTAATTGGACCTAGAAAAATATGTGGATACTCATATATTACCGCAGAA	2929

QY	901	CAATTCCTTAACGATCCATATTTAATAGAATCACTACTGTAAGCCGCTACAAAATAACAAT	960
DB	2930	CAATTCCTTAACGATCCATATTTAATAGAATCACTACTGTAAGCCGCTACAAAATAACAAT	2989
QY	961	CTTTCGATTTACATCAAAAGGAATATATCTGTTGGATTGGAGATTAGAACATTTTCAAAAA	1020
DB	2990	CTTTCGATTTACATCAAAAGGAATATATCTGTTGGATTGGAGATTAGAACATTTTCAAAAA	3049
QY	1021	ACCGATAATCTACGTCATATGTTGATTCCTCGTTTCTGTTATTTTGGTGGGTAATGTTGCA	1080
DB	3050	ACCGATAATCTACGTCATATGTTGATTCCTCGTTTCTGTTATTTTAGTTGGGTAATGTTGCA	3109
QY	1081	TTTTCTAAAGAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTTAATCATTTGGGG	1140
DB	3110	TTTTCTAAAGAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTTAATCATTTGGGG	3169
QY	1141	GGCGAAGATGTAGAATTTGGTTACAGATTATTTGCCAAAGCTGTTTTTTCAGAGTAATTT	1200
DB	3170	GGCGAAGATGTAGAATTTGGTTACAGATTATTTGCCAAAGCTGTTTTTTCAGAGTAATTT	3229
QY	1201	GACGGCGAATGGCCATCCATCAAGAACCCACTCTGTAAAGAAAAATGAAACAGACGCGAA	1260
DB	3230	GACGGCGAATGGCCATCCATCAAGAACCCACTCTGTAAAGAAAAATGAAACAGACGCGAA	3289
QY	1261	GCTGGTAAAAAGTATTACGCTTAAATTTGTGAAGAAAAAGGTACCTTACATCTATAGAAG	1320
DB	3290	GCTGGTAAAAAGTATTACGCTTAAATTTGTGAAGAAAAAGGTACCTTACATCTATAGAAG	3349
QY	1321	CTTTTACCAATAGAAGATTACATATTTCATAGAATACCTTTAGTTTCTTATTTATATCCCC	1380
DB	3350	CTTTTACCAATAGAAGATTACATATTTCATAGAATACCTTTAGTTTCTTATTTATATCCCC	3409
QY	1381	GCTTATAACTGTGCAAAATATATTCAAGATGTGTAGATAGTCTCTTTAATCAAACTGTT	1440
DB	3410	GCTTATAACTGTGCAAAATATATTCAAGATGTGTAGATAGTCTCTTTAATCAAACTGTT	3469
QY	1441	GTGCATCTCGAGGTTGTTGTTTAAACGATGGTTCAACAGATAATACCTTAAAGATGATC	1500
DB	3470	GTGCATCTCGAGGTTGTTGTTTAAACGATGGTTCAACAGATAATACCTTAAAGATGATC	3529
QY	1501	AATAAGCTTTATGTTAATATCTTAGGTTAGGTCATGCTTAAACCAAAATGGCGGAATA	1560
DB	3530	AATAAGCTTTATGTTAATATCTTAGGTTAGGTCATGCTTAAACCAAAATGGCGGAATA	3589
QY	1561	GCCTCAGCATCAAAATGCAGCGGTTCTTTTCTTAAAGGTTATTACATTTGGCAGTTAGAT	1620
DB	3590	GCCTCAGCATCAAAATGCAGCGGTTCTTTTCTTAAAGGTTATTACATTTGGCAGTTAGAT	3649
QY	1621	TCAGATGATTTATCTTGAGCCCTGATGTCAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT	1680
DB	3650	TCAGATGATTTATCTTGAGCCCTGATGTCAGTTGAACTGTGTTTAAAGAAATTTTAAAGAT	3709
QY	1681	AAAACGCTAGCTTCTGTTTATACCACTAATAGAAACGTCATCCGGATGGTAGCTTAATC	1740
DB	3710	AAAACGCTAGCTTCTGTTTATACCACTAATAGAAACGTCATCCGGATGGTAGCTTAATC	3769
QY	1741	GCTAATGCTTACAATTGGCCAGAAATTTTCAGAGAAAACTCACAACGGCTATGATGCT	1800
DB	3770	GCTAATGCTTACAATTGGCCAGAAATTTTCAGAGAAAACTCACAACGGCTATGATGCT	3829
QY	1801	CACCATTTTGAATTTTACGATTAGAGCTTTGGCATTTTAAACGGATGGATTTTAAACGAAAT	1860
DB	3830	CACCATTTTGAATTTTACGATTAGAGCTTTGGCATTTTAAACGGATGGATTTTAAACGAAAT	3889
QY	1861	ATTGAAACCGCGTGGATTATGACATGTTTCTTTAACTCAGTGAAGTTGGAAAAATTTAA	1920
DB	3890	ATTGAAACCGCGTGGATTATGACATGTTTCTTTAACTCAGTGAAGTTGGAAAAATTTAA	3949
QY	1921	CATCTTAAATAAATCTGCTATAACCGCTATTACATGTTGATACACATCCATTAAGAAA	1980
DB	3950	CATCTTAAATAAATCTGTTTATACCGCGTATTACATGTTGATACACATCAATCAATTAAT	4009
QY	1981	CTCGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGCAAGGC	2040

Db 4010 CTAGACACTCAAGAGAAATCAATTCGTTGCGTAAATCAATCATTAATAGCGCAGAGA 4069
Qy 2041 ATCAATATTATAATATGACAAATTTGATGATTTAGTAGTAAAGTAAAGATATATCTTC 2100
Db 4070 GTAAGTAATATAACTATGATGAATTCGATAACTAGATAGTAAAGTAAAGATATATCTTC 4129
Qy 2101 AATAAAACGGTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
Db 4130 AATAAAACGAGATATATCAAGAGAAATGGATATGTTTAAAGATCTTAAAGATATATCTTC 4189
Qy 2161 AATAAGATGCAAAATCGCAGTCAATTTCTTATCCCAATACATATTAACCGGCTTAGTG 2220
Db 4190 CGTAAGATGCAAAAGTGCATATAGTATATTTTATCCAAATAGATAGATGGGTAGTT 4249
Qy 2221 AAAAACTAAACAATATTATTGAATATATAAATAATATATTCGTTATTTATTCATGTTT 2280
Db 4250 AAAAAATTAAACAATATTATCGAATATATAAATAATGTTACTCATTTATTTTACACATT 4309
Qy 2281 GATAGAATCATCTTACACCACATCAAAAGAAATATTGGCTTCTTATCATAGCAC 2340
Db 4310 GACAAAATCATCTCACTTACAGACATCAAAAGAAATATCTAGAATTTCTATACAGAAT 4369
Qy 2341 CAAGTGAATATTTACTAATAATGACATCTCATATTTACACAGTAAATAGACTTAATAAA 2400
Db 4370 CAATCAATATCTGTTAAATATGATGTTCTTATATACAAACACAGGCTCATATAAA 4429
Qy 2401 ACTGAGGCACATTTAAGTAAATTAATAAATTAAGTCAGTTAAATCTAAATTTGTGAATC 2460
Db 4430 ACTAAAGCTCATTTAAGCAATATGAACAAAGTTAAGCAAACTAAATCTTAATTTAGAAT 4489
Qy 2461 ATCATTTTGTGAATCATGACAGCCTATTCGTTAAATAAGTACAGCTATGTTTATATGAAA 2520
Db 4490 ATTATTTTGAACACATGACAGCCTATTTTAAATAATGATAGTTTATACCATATATAAA 4549
Qy 2521 AATATGATCTCGGCATGAATTTCTCAGCATTAACACATGATTTGGATCGAGAAAATCAAT 2580
Db 4550 AATATGATCTCGCATGAATTTCTCCTCATTAACAAATGATTGGATTAATAAATAATTAAT 4609
Qy 2581 GCGCATCCACATTTAAAGAGTGATTAACACCTATTTTAAATCAATGACTTTAAGAAGT 2640
Db 4610 GCACACTCACCATTTAAATACTGATAAATAATATTTTAAATGATATGATTTTAAATAA 4669
Qy 2641 ATGAATGTGAAGGGCATCACAGGTATGTTTATGAAGTATGCGGTACCGCATGAGCTT 2700
Db 4670 ATAAATGAAGGCGGCTCACAGGAATGTTTATGAATATACATTTAGCACATGATATT 4729
Qy 2701 CTGACGATTTAAAGAGTCATCACATCTGCCAATCAATTTGATAGTGCAGAGATAT 2760
Db 4730 GCTACGATTTAAGAGAGTCATATACATTTATGCCAATCTACAGATAGTGTCTGTAAT 4789
Qy 2761 AACACTGAGATATTGGTTCCAAATTTGCACTTTTAAATCTTAGAAAAAGAAACCGGCAT 2820
Db 4790 AATCTGAAGATATTTGGTTCCAAATTTGCACTTTTAAATCTTAGAAAAAGAAACCGGCAT 4849
Qy 2821 GTATTTAATAAACATCGACCTGACATTTATATGCTTTGGGAACGAAATTTACATGACA 2880
Db 4850 GTATTTAATAAACATCGACCTGACATTTATATGCTTTGGGAACGAAATTTACATGACA 4909
Qy 2881 AATGAACAATTTCAAGTGCAAAAAGGGAATAATATCCCGTTTACAAGTTTCAATTATT 2940
Db 4910 AATGAACAATTTGAAGTGCAAAAAGGGAATAATATCTGTTTAAACAGTTTCAATTATT 4969
Qy 2941 AATAGTATAAGCTTATAAACATTTGCAATTTTATTAATA 2979
Db 4970 AATAGTATAACTCTATAAAACATTTGCAATTTTATTAATA 5008

RESULT 3
LOCUS AF302467 8838 bp DNA linear BCT 12-MAR-2001
DEFINITION Pasteurella multocida P4218 region 2 capsule biosynthesis gene cluster, partial sequence.

AF302467
AF302467.1 GI:13274373
pasteurella multocida.
pasteurella multocida.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
pasteurella.
1 (bases 1 to 8838)
Townsend, K.M., Boyce, J.D., Chung, J.Y., Frost, A.J. and Adler, B.
Genetic organization of Pasteurella multocida cap Loci and
development of a multiplex capsular PCR typing system
J. Clin. Microbiol. 39 (3), 924-929 (2001)
21142635
11230405
2 (bases 1 to 8838)
Townsend, K.M., Boyce, J.D., Chung, J.Y., Frost, A.J. and Adler, B.
Direct Submission
Submitted (03-SEP-2000) Veterinary Pathology and Anatomy, The
University of Queensland, School of Veterinary Science, Brisbane,
QLD 4073, Australia
Location/Qualifiers
1. 8838
/organism="Pasteurella multocida"
/strain="P4218"
/serotype="F:3"
/db_xref="taxon:747"
complement(<1. .351)
/gene="phyA"
complement(<1. .351)
/gene="phyA"
/note="similar to KpsS, LipA"
/codon_start=1
/transl_table=11
/product="PhyA"
/protein_id="AAK17919.1"
/db_xref="GI:13274374"
/translation="MSILYDIRLNMYOEFLIFSGKMLKIPYLSGFFTORLKMFSFV
TWKERTCILEWGYKASSRKARHFAQQHDLPTATIEDGLRSIGLGVDPYPPSLYVD
DIGIYDINQPSRL"
699. .2567
/gene="fcbE"
699. .2567
/gene="fcbE"
/note="similar to Hyae, KfiB"
/codon_start=1
/transl_table=11
/product="FcbE"
/protein_id="AAK17920.1"
/db_xref="GI:13274375"
/translation="MKVILIGHKQSNQDYKVFQCYGNPNPLPSKREKNSPIEIGH
VLMKVLPSUEHTPKNSLLSNKSKIKKNSAKNSHKHAKTNTIOTTSSIMWNLSD
LMANIEQNFQWSDPNAIQILDYWANLDPNIHFVYDKPENLFQVHSLLEALKLDK
HTVQEFQEWQTYNEKILTYFNKYKDRSVLLNTQOOLNTKRTSLSEIYKHLSAPDALV
YISEQILVNDKEMEIIEVNQDLSHQEECPNSFVSIKNSPTVQYVEELQSHADLP
KLNEPFLNAMEDIFALLAKMDIQKVNOYQHEKELELSTIKERQLEVTERVQLTEQ
KLSETOKELOIKDENRKVKSEKAKLTASVOSTSKILSEKEKETSICLSKENTKIEEK
IKDEAYHLTKTSLDSKALKTHODEIPEALKIENENISVQEDMQEKFQENKROE
LEQELKASDKKALLETENSKQTSVESLENENKVLAAQLOLQOELEKLYIDNOVLK
AKPLRYGAADRIKQNTYRGYKIQHRSILGFLIFLPIFFLYLGLFKREKMKRYEN
TLPPIHIEYDEAHEANRISKHLSKYLGLVFLQEIINNPNPKWLTLPYKLLKEGRFKQG"
2581. .5481
/gene="fcbD"
2581. .5481
/gene="fcbD"
/note="similar to PmcS"
/codon_start=1
/transl_table=11
/product="glycosyltransferase FcbD"
/protein_id="AAK17921.1"
/db_xref="GI:13274376"
/translation="MNTLSQAIKAYNSNDYELALKLFEKSAETVGRKIVFQIICKKE
KLSTNSVSEDKNSVCDSDSIATQLLLSNVKKLTLSSEKNSLKNKKNSITGCKKSE
NAEIRKVELVPKDFPKDLVLAFLPDHVNDFTWYKRRKSLGIRKPNKNNIGLSITPTF

NRSRLDITSAFLVNQKTYNPREVVVADDSKSNLLTIVQVYEOKLDIKYVROKDYG
OLCAVNLGLRLTAKYDFVSLDDMAPOQLWVHSLYTELLEDNDLVLIGPRXYVDTHN
TIAEOLFNDPYLESLPETATNDNPSITSKGNISLDWRLHLEFKKTDNLRCLDPSRPFY
SCGNVAFKSEWLNKLGWDFEENHGGEDVEFGYRLFAGCFFRVIDGMAVHQBPFG
KENETDRAGKSLIKVIREKVPYIYRKLLPDEDSHRIPLVSIYPAYNCANYIOR
CVDSALNVTVDLEVICNDGSDNTLEVINLGNPNRVRIMSKPNGLIASNAVA
SPAKGYIGOLDSDYLDLPDAVELCLKEFLKDKTLACVYTNRNVPDGLSLANGVW
PEFSREKLTAMIAHFERMTIRAWHLTDGFNEKTENAVVDYDMFLKLSVGFREKHLN
TCYRVLHGDNFTSIKNDLTKQKNEHVYNQSLNRQVSNYNDEFDNLDSEKYEFNK
TADYQEDILADIKIVQRDAKVAISIPNRLDGLVKRLNNIIEYNKNVLIVLHI
DKNHLTSDIKLEIFBEHNKNOINILLNDVSYTNNRLIKTKAHLNSNMKNLROLNLI
EYIIFDNHDSLPFKDSYNIHKYDIGNFSSLTNDWINKINAHSPFKNLIRKYFNDN
DLKTNMKGASQMGKIKYTLAHDIAIATINKEVITLCQSTDSVPEYNTEDWFOFALLIL
EKTCHVENKSTSTLYMPWERKLTQNEQIESAKRGENIPVKNFIINSITL
5545..6717
/gene="fcbC"
5545..6717
/gene="fcbC"
/codon_start=1
/transl_table=11
/product="UDP-glucose dehydrogenase FcbC"
/protein_id="AAK17922.1"
/db_xref="GI:13274377"
/translation="MKKTIAGAGYVGLSNVLAQHNNVILLIDQNKVDLINNKS
PITDKMEDFLQNKSTMATPDKEVALKNADVFVNATPDYNTETGYENTSTVEAVI
EOTLSINPOAAIIIKSTIPVGFTENNREKFNPLIFSPFLREGKALVDNLYPSRII
VGSSTYQAKVADMLTQCAKEDVTLFTHNTEAEAVKLFANTYLAARVAFNPLDYY
ASLHLNKTIDINGISTDRGTHNPNFSGYGGYCLPKDQKQLLANVADVPQNLIEA
IVKSHNETRRFTHDVLNKNKPTVGIYRLIMKSGSDNFRASAILDIMPHEKNGVEIV
IYKPTLNQOAFEDYFVNIQLSEFINRSQVILANRSEPDNLQCSHKIYTRDIFGGDA"
6710..8143
/gene="fcbB"
6710..8143
/gene="fcbB"
/codon_start=1
/transl_table=11
/product="glycosyltransferase FcbB"
/protein_id="AAK17923.1"
/db_xref="GI:13274378"
/translation="MLNFKIHKYVHIQSLHLKKEYALLYAKYINQLSINOQAYVIC
QULKLDLIDPKWHSVFFQLGILARGHDSDEMRRLITCTDFSKNKQILSOLL
AYSQIATLCPQTYRYALYLSLANLKDLYFLKEELNKLPLSCVLNPHYCLLQNF
VEKNSKIKENINQFLYFKLGEITITKEDPPNTNIVSHSVNDNTYLNAPLITL
LVTFNSQKSLNLTNSLFNQSPNIEIIVIDHSDNDMSILOAYTKQYKNKITSL
KENQVTVAKNIGLYKASGEFTLQCDSDHMAHQKALQVAPLQKELLIVYFESKVR
LDPIGNYPARTIYPLMRNPSPALFKKEVCEKTLADWVRIGADSEFNARLKLIFGH
KGYTVNKPPLTGAHRENSLMTAQSTGYVNGVSLPREAYKAWNIWHISOLQORHAPL
LSSSPKREFQYDQHPIMDVISYAIIN"
8162..>8838
/gene="hexD"
8162..>8838
/gene="hexD"
/note="similar to BexD"
/codon_start=1
/transl_table=11
/product="HexD"
/protein_id="AAK17924.1"
/db_xref="GI:13274379"
/translation="MKPIKTLISFTITLILAGCHSMPTSGPQNHIIGLKPKQNSL
PSVDVTEMDNKVAHTLFLKQKQSQSTQFKQNSYADIINVGDITDLVLIWEAAPALIF
GSVLQSTGGGANLTLPQIVARNGKITPFLGPLVKGKTPQEIORDIWAHALSSLA
NKQVTVRLNKNKSNKVTILRQNSVRMPLTSGQERVDAIAAVGGATENLQDISVOL
TRGEVK"

gene

CDS

gene

CDS

gene

CDS

BASE COUNT 3387 a 1519 c 1360 g 2572 t
ORIGIN

Query Match 88.4%; Score 2632.6; DB 1; Length 8838;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 209; Indels 1; Caps 1;
QY 1 TTATAACTGATTAAAGAAGGTAACAGGTTCAAGCAAGGTTAAATTTTAAAGGAAGAAA 60
DB 2525 TTATAACTGATTAAAGAAGGTAACAGGTTCAAGCAAGGTTAAATTTT -AAGGAAGAAA 2583

QY 61 ATCAATACATTATCAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
DB 2584 ATGAATACATTATCAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC 2643
QY 121 AAATATTTTGAAGTCTGCTGAAACCTACGGGGGAAAAATCGTTGAATTTCCAAATTTATC 180
DB 2644 AAATATTTTGAAGTCTGCTGAAACCTACGGGGGAAAAATCGTTGAATTTCCAAATTTATC 2703
QY 181 AAATGTAAGAAAAAATCTCTCGCAATATCTTATGTAGTGAAGATAAAAAACAGTGT 240
DB 2704 AAATGTAAGAAAAAATCTCTCGCAATATCTTATGTAGTGAAGATAAAAAACAGTGT 2763
QY 241 TGGGATAGCTCATTAGATATCGCAACACACCTCTTACTTTCCAAACGTAAAAAAATTAAC 300
DB 2764 TGGGATAGCTCATTAGATATCGCAACACACCTCTTACTTTCCAAACGTAAAAAAATTAAC 2823
QY 301 CTATCCGAATCAGAAAAAATACAGTTTAAAAAATAAATGGAATCTATCTACTGGGAAAAA 360
DB 2824 CTATCCGAATCAGAAAAAATACAGTTTAAAAAATAAATGGAATCTATCTACTGGGAAAAA 2883
QY 361 TCGGAGAACCGCAAAATCAGAAAGGTGGAAGTAGTACCCCAAGATTTTCTTAAGATCTT 420
DB 2884 TCGGAGAACCGCAAAATCAGAAAGGTGGAAGTAGTACCCCAAGATTTTCTTAAGATCTT 2943
QY 421 GTTCTTCTCCATTGGCCAGATCATGTTAATGATTTTACATGGGTACAAAAATCGAAAAAAA 480
DB 2944 GTTCTTCTCCATTGGCCAGATCATGTTAATGATTTTACATGGGTACAAAAATCGAAAAAAA 3003
QY 481 AGCTTAGGTATAAGCCCTGTAATAAGATATCGGTCTTTCTATTATTATTTCTTACATTT 540
DB 3004 AGCTTAGGTATAAGCCCTGTAATAAGATATCGGTCTTTCTATTATTATTTCTTACATTT 3063
QY 541 AATCGTAGCCGTATTTTAGATATATACCGTTAGCCCTGTTTGGTCAATCAGAAAAAATAC 600
DB 3064 AATCGTAGCCGTATTTTAGATATATACCGTTAGCCCTGTTTGGTCAATCAGAAAAAATAC 3123
QY 601 CCATTTCAAGTCGTTGTTGTCAGATGATGTTAGTAAAGAAAACTTACTTACCATTGTGCAA 660
DB 3124 CCATTTCAAGTCGTTGTTGTCAGATGATGTTAGTAAAGAAAACTTACTTACCATTGTGCAA 3183
QY 661 AAATACGAACAAAAAATTTGACATAAAGTATGTAAGACAAAAAAGATTATGGATATCAATTG 720
DB 3184 AAATACGAACAAAAAATTTGACATAAAGTATGTAAGACAAAAAAGATTATGGATATCAATTG 3243
QY 721 TGTGCACTCAGAAACTTAGGTTTACGTACAGCAAGTATGATTTTGTCTCGATTCTAGAC 780
DB 3244 TGTGCACTCAGAAACTTAGGTTTACGTACAGCAAGTATGATTTTGTCTCGATTCTAGAC 3303
QY 781 TGGGATATGCAACCAACAAATATGTTGTTTCAATTTCTTATCTTACAGAACTATTAGAAGAC 840
DB 3304 TGGGATATGCAACCAACAAATATGTTGTTTCAATTTCTTATCTTACAGAACTATTAGAAGAC 3363
QY 841 AATGATATTTGTTTAAATTTGGACCTAGAAAAATATGTTGATCTACTATAATTTACCGCGAAA 900
DB 3364 AATGATATTTGTTTAAATTTGGACCTAGAAAAATATGTTGATCTACTATAATTTACCGCGAAA 3423
QY 901 CAATTCCTTACGATCCATATTTAATAGAATCACTACTGAAACCGCTACAATAAACAAT 960
DB 3424 CAATTCCTTACGATCCATATTTAATAGAATCACTACTGAAACCGCTACAATAAACAAT 3483
QY 961 CTTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGAACATTTTCAAAAAA 1020
DB 3484 CTTTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGAACATTTTCAAAAAA 3543
QY 1021 ACCGATAATCTACGCTCTATGTGATTTCCCGTTTCCGTTTATTTTGGGGGTAATGTTGCA 1080
DB 3544 ACCGATAATCTACGCTCTATGTGATTTCCCGTTTCCGTTTATTTAGTTCGGGTAAATGTTGCA 3603
QY 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGGTTGGTTGCGATGAAGAAATTTAATCATTTGGGGG 1140
DB 3604 TTTTCTAAAGAAATGGCTAAATAAAGTAGGTTGGTTGCGATGAAGAAATTTAATCATTTGGGGG 3663

JOURNAL Submitted (19-MAY-1998) Microbiology, Monash University, Wellington
Road, Clayton, VIC 3168, Australia
REFERENCE 3 (bases 1 to 16727)
AUTHORS Chung, J.Y., Zhang, Y. and Adler, B.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2000) Microbiology, Monash University, Wellington
Road, Clayton, VIC 3168, Australia
REMARK Sequence update by submitter
COMMENT On May 5, 2000 this sequence version replaced gi:3435183.
FEATURES
source
1..16727
/organism="Pasteurella multocida"
/db_xref="taxon:747"
complement(1157..2383)
/gene="phyB"
complement(1157..2383)
/gene="phyB"
/note="LipB, kpsS homolog"
/codon_start=1
/transl_table=11
/product="PhyB"
/protein_id="AAC67247.1"
/db_xref="GI:3435184"
/translation="MDVLKHNLDLLESSNYYLLQGPIGPFFTELSAWLTKHQQVF
KINFSGDEAFYNSIPQTFAYQGLDQDYLRFFCHQHSIDAIICFGDTRAYHRIA
KIVADQKISFWFEFGYLRPEYVTLKGGVNDYSQLPHANFELLSQAACCSISAPE
KLAKGFIPLATIAIKYTKAYFNQDKYPDYQHURILDLYIKLWLTSGIKRAYVYQW
DRSFGKKVAGKGLDFELVPLQVDDSQVRIHCDYESVDAFLKELLESFTHAPSALN
LIKHHPMDRGFYDGHVIGKQYQTPAFKRIFYVHDVPTILLRKGKGNITLNSYS
GIALLRMPVLVLGRASYNFEGTLTYGNNLHSEFHNKGLPEAKVFAYRYKYLKSTQI
NGSFYSKVLIRYPYQ"
complement(2393..4483)
/gene="PhyA"
complement(2393..4483)
/gene="PhyA"
/note="KpsS, LipA homolog"
/codon_start=1
/transl_table=11
/product="PhyA"
/protein_id="AAC67248.1"
/db_xref="GI:3435185"
/translation="MSTLYDYIRLNMVQEPILFSKGLMKIPIYLSGFFTORLKMFSPEV
TWKRTCLIEWGYSKSKRAHFAQHDLPYATIEDFLRSIGSLGVDGYPFPLSLYVD
DIGIYDINQPSLENLILSQDLQAEKQVQVEYATIELICTHNLKYNHAIPTPLQNT
KRPIVLVDQYTGMAVTGNAEQSDFLHMLERAIITENPTAEILTKLTHDPDMCGKQOG
YLTXEQPFRVKVLSDEPNSISLLKHVDKVCYVTSHTGFEALLLGTIVYTFGAAWESG
WGLTDRHAYIRQKSKRAKRSLLQQLFYAATFYQCRYINPNTGKSGTDFDVIDLI
QAKKVTNQLAGDIYCGMRPWRKRVQPPFPFPRCLHFLVNLHKLRCITHEKAQAKI
VYMGHSHIEVYEAQQQLPRLRMEDGFLRSVGLGSLNLTPTISLVLDVDVGIIYFDAQSR
SRLEDILQHSFTLKLQRAETLTKTLIEQHIGKYNVGHPTHLCLTHRONKLLVVGQV
ENDASIQGSPHIRTNAELLCVTRKNNPOAYIYKPHDPVAVGNKNTDRLLDYROYA
DFVVEKVNILDCINQDVEVHTMTSLAGFEALLREKKVHCYGLFPYSNWGLTVDHLSLN
RRSRKLSLELIAGVLIYPOYIDPKTKTMIDVQRAVDIILIEQRRKIKNKNKLNTYFM
NIFMKLKNVSVLR"
4832..6700
/gene="hvae"
4832..6700
/gene="hvae"
/note="KfiB homolog"
/codon_start=1
/transl_table=11
/product="Hvae"
/protein_id="AAC67249.1"
/db_xref="GI:3435186"
/translation="MKKVIIGHKQSNQVQVEKVFQCYGMNPLPDKREKMSPIETGH
VLNKLPSFEHTPKNYSLLSKSIKKGSAKNKSHKAKTNTIQTTSINWNLSD
LMLANTEQWGSNDPNAIQILDYWANLPNIHFVYVDYKPNFLFYHSLSEALKLDK
HTVQKEFEWQTYNEKILTYFNKYKDRSVLLNTQQLQNTKTKTSILSYKHISAPDALY
KLINEPSLNKEMEIIEVNODLSHOEECPLSNFIVSQIKNPSPTVQVYRELOSHADLP
YISEQKLVNDADALLAWKDMIOKQVANDQYHKEKELELSTIKERQLEVTEYKQLTEQ
KLSEQKLEIQIKDENRKYKSEKALTASVQSTKILSEKEKESICKSSENTKIKEK
IKIDEAYHTKTLSDKALKTHQDEIEALKIIFNENTSVQEDMQEKFOETNKKQOE
LEQELKAIKSDKALLETENSQKTQVSESLENENKVLLAQQLIQEIELEKLYIDNQVLK

gene
6717..9635
/gene="hyad"
6717..9635
/gene="hyad"
/note="Hyad"
/codon_start=1
/transl_table=11
/product="glycosyl transferase"
/protein_id="AAC67250.2"
/db_xref="GI:7710190"
/translation="MNTLSQAIKAYNSNDYELALKFEKSAEYIGRKIVEIFQITCKRE
KLSAHPSVNSAHLVSVNKEKVCNCDSPDIDATOLLNVKKLVSLSSEKNTKPKHQQVGL
LTEKKSSEAVRAVALVPKDPKDLVLAPLPDHVNDFTWYKKRKLGIKPEHQHVGL
SIIVTFNRPAIISITLACLINQKTHYFVFIIVDDGSEDLSPIIRQYENKLDIRYV
KQKNGFQSAARNMGLURAKYDFIGLDDCDMAPNPLVHVSIVAEELLEDLDTIIGPR
KYIDQHDIDPKFLNNASLESLEPVTNNNSAAKGGTGVSLMDRLQFQEFKTLNLRIS
DSPRFFPAAGNAFAKAKNLKSGFFDEEFNHMGGEVFGYFLFRYSGSFFKTIQDMA
YHOEPCKENETDREACKNTLDMREKVPYIRKLLP IEDSHINRVPLVSIYIPAYN
CANYIORCVDSALNOTVVDLEVCICNDGSTDNTLEVINKLYGNPRVIRINSKPNGGIA
SASNAVSFAKGYIIGQLDSDDDILEPDAVELCLKEFLKDKTLACVYTNRNVPDGLS
IANGYNMPEFSREKLTAMIAHFMETIRAWHLTDGFKENIENAVDIDMFLKLSVGL
KFHLNKCINRVLRHGDNTSIIKLGIOKKNHVVYVNSLNRRQGTYYNYDEFDLDES
RKYIFNKATYQEEIDILDKIITIQNKDAKIAVFIYPNTLNGVLKLNNTIENYKNI
FVIVLHVKNHLTPDIKKELAFYHQHVNLINLANDISYVTSNRLIKTEAHLIENKIL
SOLNLANCEYIIFDNHDSLFVKNDSYAMKYDVGWMSFALTDHWEIKINAHPPFKLLI
KTYPNNDLSMNVKGASQGMFTYALAHALLTTIKEVITSCOSIDSVPYENTEDIWF
QFALLILLEKTKGVFNKTSITLTYMPWERKQLQMTWEQIESAKRGENIPVNRKIINSITL

gene
9699..10871
/gene="hyaC"
9699..10871
/gene="hyaC"
/note="hyaC"
/codon_start=1
/transl_table=11
/product="UDP-glucose dehydrogenase"
/protein_id="AAC67251.2"
/db_xref="GI:7710191"
/translation="MKKITTAGYVGLSNVLAQHNNVILLDQNKVDLINKKKS
PITDKEJEDLQNKSLMTMATDKVAKLNADFVIAITPTDYNTEGYFNTSTVEAVI
EQGTSINQPAITIIKSTIPVGTENREKFTNPLIFSPFVREGKALYDNLVPSRII
VQTSYQAKVAFADMLTQCARKQDVTLFTHTTEAEAVKLFANTYLANRVAFNEDLTY
ASLHLNKTDIINGISDPRIGTHYNNPSCYGCYCLPKDTKOLLANVADVPQNLIEA
IVKSNTRKRFIHDVNLNKPXTVGIYIRLIMKSGSDNFRASAILDIMPFLKENGVEIV
IYEPFLNQAFEDVPVNLQSEFINRSVDVILANKSEPDNLNOCSHKIYTRDIFGGDA"
10864..12297
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="Hyab/Hyaa fusion protein"
/protein_id="AAC67252.2"
/db_xref="GI:7710192"
/translation="MLNLFKIKRYVHHQSLHKKKEYALLVAKYINLLISNQQAAYVIC
OLKLYDLFLDPKWSHVSFFOLGLIARGHADNDVMVRLITCTDFSKNKLILSOLL
AVSQIATTLDPKTYRYRALLYSLVANKDPRVLKEELNKLPSVLNTPHYVCLLOFI
VERENSKLENINQFLFYKLGTEITTKTEPPNTNIVSHVNDNTYINLAPLITI
LVTFNSQKSIKNTLSNFQSYNPIEIVIDHSDQNTMSILQAYTKQYKNIKIISL
KENYGVYAKNIGLKYASGEFITCODSDWAHPQKLALQVAPLLOHRELIVTFSKWVR
LDPIGNPYARTIYPLMRNLPSALFRKKEVCEKTAALMDWVRIGADSEFNARLKLIFGH
KGYVTYVNRKPLTFGAHRENSLMTAQSTGYVNCVSLPREAYKNKAWNIWHISQLOQRHAPL
LSSSPKREFQYDQHPIMDVISYAIN"
12316..13497
/gene="hexD"
12316..13497
/gene="hexD"
/note="CpxD, BexD homolog"
/codon_start=1
/transl_table=11
/product="HexD"
/protein_id="AAC67254.1"
/db_xref="GI:3435191"

gene
CDS
/translation="MKPIKTLISFITTLLIAGCHSMPTSGPAONHILIGLKKPONESL
PSVDIEMNDKVAITLFKQKOSQFTOKOONSNYADIINVGDITLDVLWEAPAILF
GSLVSQVSGSGLNLTPEQIVARNKGIITPFILPKVILVKTPEQIORDLAHALSSLA
NKPOVIVRLNNKSNVTILRQGSVRPLTSQGERVLDALAAVGGATENLDLSVOL
TRGEVKLSLEKLATHPEENILRNSDVVTLLKPLSFTGLGALGNTKOVKFSANGL
TLABGICEMGLLDNRADPKGVFVRIIPFNKLSLAQETKWKARGYDNDEIPTVSV
LLNPNALFWLQRPFIQDKDLVYVSNAPMAEFQKFLKLVFSITSPVTGTLHNINIKN
L"
13507..14643
/gene="hexC"
13507..14643
/gene="hexC"
/note="hexC, CpxC homolog;
MIEFKIQKNOFKKWSRFRKISTLFKVMVILPTCCSLFYFSLWASDIYISQSPFVVT
PKNOALSGVALLOSGFARAQDDTYVQEFMRSTRLELLEKSIPIRQFVEDKGD
FSRPNLISFEQAFQYFSKLSLVNFDVSGIATLINIRAFDREAQIQNOELLKOG
EYLINRNEARKDTVMFAELVASEAKKVTETSSALSEYKRNKGVFDLOSQSEVLS
LISTQORLNIQTOLDQVRSISPNPQVOTLLARANSIRKEMOQVOVLGGNSIV
TOTAEYORLVNNTLAQQLGTATISLQNRSEADROQLYLEIISYPNEPDLAEPYR
LYNTLATFLSILYGITLLLASIREHN"
/codon_start=1
/transl_table=11
/product="HexC"
/protein_id="AAC67255.1"
/db_xref="gi:3435192"
/translation="MIEFKIQKNOFKKWSRFRKISTLFKVMVILPTCCSLFYFSLWA
SDIYISQSPFVVTPKNOALSGVALLOSGFARAQDDTYVQEFMRSTRLELLEK
SIPIRQFVEDKGDLSFRPNLISFEQAFQYFSKLSLVNFDVSGIATLINIRAFD
KEAQIQNOELLKOGEYLINRNEARKDTVMFAELVASEAKKVTETSSALSEYKRN
GVFDLOSQSEVLSLISTQORLNIQTOLDQVRSISPNPQVOTLLARANSIRKEMO
QVOVLGGNSIVTOTAEYORLVNNTLAQQLGTATISLQNRSEADROQLYLEIIS
YPNEPDLAEPYRLYNTLATFLSILYGITLLLASIREHN"
14643..15440
/gene="hexB"
14643..15440

Query Match 78.6%; Score 2340.8; DB 1; Length 16727;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 372; Indels 22; Gaps 2;
QY 1 TTATAACTGATTAAAGAAGGTAACGATTCAAGCAGGTTAAATTTTAAAGGAAGAAA 60
DB 6658 TTATAACTGATTAAAGAAGGTAACGATTCAAGCAGGTTAAATTTT-AAAGGAAGAAA 6716
QY 61 ATGAATACATTTATCAACAGCAATAAAGCATATACAGCAATGACTATGATTAATAGCACATC 120
DB 6717 ATGAATACATTTATCAACAGCAATAAAGCATATACAGCAATGACTATGATTAATAGCACATC 6776
QY 121 AAATTATTTGAGAAGTCTGCTGAAACCTACGGCGGAAAAAATCGTTGAATTCCAAATATATC 180
DB 6777 AAATTATTTGAGAAGTCTGCGGAAATCTATGGACGGAATAATTTGTTGAATTTCAAATATACC 6836
QY 181 AAATGTAAGAAAACTCTCGACCAATTC-----TTATGTTAACT 219
DB 6837 AAATGTAAGAAAACTCTCAGCACATCTCTGTTAAATTCAGCACATCTTTCTGTGTAAT 6896
QY 220 GAAGTAAAAAACAAGTCTTTCGATAGTCTCATATAGATATCGCAACACAGCTCTTACTTT 279
DB 6897 AAAGAAGAAAAAGTCAATGTTTTCGATAGTCTCGTTAGATATTTGCAACACAACTGTACTTT 6956
QY 280 TCCAACGTAATAAAATTAATCTATCCGAATCAGAAAAAACAAGTTTAAAAAATAAATGG 339
DB 6957 TCCAACGTAATAAAATTAATCTATCTGACTCGGAAAAAACAAGTTTAAAAAATAAATGG 7016
QY 340 AAATCTATCACTGGGAAAAAATTCGAGAGACGAGCAATCAGAAAGTGGNACTAGTACCC 399
DB 7017 AAATCTCACTGAGAGAAATCTGAAAAATGCGGAGTGAAGACGGTCCGCCCTTGTAACA 7076
QY 400 AAAGATTTTCCCTAAACATCTGTGTTCTGCTCCATTCGCCAGATCATCTTAATGATTTTACA 459
DB 7077 AAAGATTTTCCCAAGATCTGGTTTTAGCCCTTTTACCTGATCATGCTTAATGATTTTACA 7136
QY 460 TGTACAAAAATCGAAAAAAGCTTTAGGTATAAAGCCTGTAAATAAGGAATATCGGCTCTT 519

DB 7137 TGGTACAAAAAGCGAAGAAAAAGACTTGGCATAAAAACCTGAACATCAACATGTTGCTCTT 7196
QY 520 TCTATTATTATTCCTACATTTTAACTCGTAGCGGTATTTTAGATATATAACGTTAGCCTGTTTG 579
DB 7197 TCTATTATTCGTTTACAAACATTCATTCGAGCACCAGCAATTTTATCGATTACATTTAGCCTGTTTA 7256
QY 580 GTCATCAGAAAAAACAATACCATTTTGAAGTCGTTTGGCAGATGATGGTAGTAGAAGAA 639
DB 7257 GTAAACCAAAAAACAATATACCCGTTTGAAGTATCGTGACAGATGATGGTAGTAGAAGAA 7316
QY 640 AACTTACTTACCATTTGTGCAAAAAATACGAACAAAACTTGACATTAAGATATATAGACAA 699
DB 7317 GATCTATCACCGATCATTCGCGCAATATGAAATAAATAATGGATATTCGCTACGTCAGACAA 7376
QY 700 AAGATTATGGATATCAATTTGTCGAGTCAGCAAACTTAGGTTTACGTACACACAAAGTAT 759
DB 7377 AAGATAACGGTTTTTCAAGCCGTCGCCGTGCGAATATGGGATATAGCTTACCAAAATAT 7436
QY 760 GATTTTGTCTCGATTCTTAGACTGCGATATGGCACCACAAATATATGGTTTCACTCTTAT 819
DB 7437 GACTTTATTGGCTTACTCGACTGTGATATGGCGCAATCCATTTATGGTTTCACTCTTAT 7496
QY 820 CTTACAGAACTATTAGAGACAAATGATATTGTTTAAATTGGACCTAGAAAAATATGTGGAT 879
DB 7497 GTTCGAGAGCTATTAGAAGATGATGATTTAACAATCATTTGGTCCAAAGAAAAATACATCGAT 7556
QY 880 ACTCATATATATACCGCAGACAAATTCCTTAACGATCCATATTTAATAGAATCACTACCT 939
DB 7557 ACACAACATATATGACCCCAAAAGACTTCTTAAATACGCGAGTTTGGTTGAATCATTTACCA 7616
QY 940 GAAACCGGTACAAAAATAACAATCCTTCGATTACATCAAAAGGAAATAATATCGTTGGATGG 999
DB 7617 GAAGTGAACCAATAATATAGTGTGCCGCAAAAGGGGAAGAACAGTCTTCTCTGGATGG 7676
QY 1000 AGATTAGAACATTTCAAAAAACCGATATCTACGCTCTATGATGATTCCTCGTTCTGTTAT 1059
DB 7677 CGCTTAGAACAAATTCGAAAAACACAGAAAAATCTCCGCTTATCCGATTTCGGCTTTCCGTTT 7736
QY 1060 TTTGTTGCGGGTAATGTTTCATTTTAAAGATAGCTTAAATAAAGTAGTGTGGTTTCGAT 1119
DB 7737 TTTGCGCGGGTAATGTTTCGCTTTCGCTTAAATAAAGTAAATAAATCCGTTTCTTTGAT 7796
QY 1120 GAAGAAATTTAATCATTTGGGGGGGGAAGATGATAGATTTGGTTTACAGATTTATTTGCCAAA 1179
DB 7797 GAGGAATTCATCACTCGGGTGGAGAAGATGTGGAATTTGGATATCGCTTATTCCTGTAC 7856
QY 1180 GCGTGTGTTTTCAGAGTAAATGACCGCGGAATGGCCATCCATCAAGAACCCACCTGTTAAA 1239
DB 7857 GGTAGTTTCTTTAAAACTATTGATGGCATATGGCCCTACCATCAAGAGCCACCGGTAAA 7916
QY 1240 GAAATGAACAGACGCGGAAGCTGGTAAAGTATTTACGCTTAAATTTGTGAAAGAAAAAG 1299
DB 7917 GAAATGAACCGATCGTGAGCGGGAATAATATTTACGCTCGATATTTATGAGAGAAAAAG 7976
QY 1300 GTACCTTACATCTATAGAAGCTTTTACCAATAGAAGATTCACATATTCATAGATACCT 1359
DB 7977 GTCCCTTATATCTATAGAAAACTTTTACCAATAGAAGATTCGATATCAATAGATACCT 8036
QY 1360 TTAGTTTCTTATTTATATCCCGCTTAACTGTGCAAAATTTATATTCAAAGATGTTAGAT 1419
DB 8037 TTAGTTTCTTATTTATATCCAGCTTATTAAGTGTGCAAACTATATTCACGTTCCGTAGAT 8096
QY 1420 AGTGTCTTAACTGTTGTCGATCTCGAGGTTTGTATTTTGAACGATGTTTCAACA 1479
DB 8097 AGTGCACTGAATCAGACTGTTGTTGATCTCGAGGTTTGTATTTGTAACGATGTTTCAACA 8156
QY 1480 GATAATACCTTAGAAGTATCAATAAGCTTTATGGTAATAATCCTAGGGTACGATCATG 1539
DB 8157 GATAATACCTTAGAAGTATCAATAAGCTTTATGGTAATAATCCTAGGGTACGATCATG 8216
QY 1540 TCTAAACCAAAATGCGGGAATAGCCTCAGCATCAAAATGACGCGCTTCTTTTGTCAAGGT 1599

Db 8217 TCTAAACCAATGGCGGAATAGCCTCAGCATCAAAATGAGCCGTTTCTTTTGGCTAAAGGT 8276

Qy 1600 TATTACATTTGGGCAGTTAGATTCAGATGATTTATCTTGAGCCGTGATGCACTGGT 1659

Db 8277 TATTACATTTGGGCAGTTAGATTCAGATGATTTATCTTGAGCCGTGATGCACTGGT 8336

Qy 1660 TTTAAAGAAATTTTAAAGATATAAACGCTAGCTGTGTGTTTATACCACATAATAGAAACGTC 1719

Db 8337 TTTAAAGAAATTTTAAAGATATAAACGCTAGCTGTGTGTTTATACCACATAATAGAAACGTC 8396

Qy 1720 AATCCGGATGATGCTTAATCGCTAAATGGTTTACAAATTTGGCCAGAAATTTTCAGAGAAAA 1779

Db 8397 AATCCGGATGATGCTTAATCGCTAAATGGTTTACAAATTTGGCCAGAAATTTTCAGAGAAAA 8456

Qy 1780 CTCACAAGGCTATGATTCCTCACCATTAGAAATTTTAGAATGTTTACGATTAGAGCTTGGCATTTA 1839

Db 8457 CTCACAAGGCTATGATTCCTCACCATTAGAAATTTTAGAATGTTTACGATTAGAGCTTGGCATTTA 8516

Qy 1840 ACGGATGGATTTAAACGAAATATTGAAACGCCGTGGATTATGACATGCTTCCCTTAAACCTC 1899

Db 8517 ACTGATGGATTCANTGAAATTTGAAATTCGCGTAGACTATGACATGCTTCCCTCAACCTC 8576

Qy 1900 AGTGAAGTTGGAATTTTAAACATCTTTAATAAATCTGCTATAACCGCGTATTACATGGT 1959

Db 8577 AGTGAAGTTGGAATTTTAAACATCTTTAATAAATCTGCTATAACCGGTATTACATGGT 8636

Qy 1960 GATAACACATCCATTAGAAACTCGGCATTTCAAAAGAAAACCATTTTGTGTAGTCAAT 2019

Db 8637 GATAACACATCCATTAGAAACTTTGGCATTCAAAAGAAAACCATTTTGTGTAGTCAAT 8696

Qy 2020 CAGTCATTAAATAGACAAGGCATCAATTTATTATTAATATGACAATTTGATGATTAGAT 2079

Db 8697 CAGTCATTAAATAGACAAGGCATCAATTTATTATTAATATGACAATTTGATGATTAGAT 8756

Qy 2080 GAAAGTGAAGATATATCTTCAATAAACCCTGAATATCAAGAGAAATGGATGTGA 2139

Db 8757 GAAAGTGAAGATATATCTTCAATAAACCCTGAATATCAAGAGAGATTTGATATCTTA 8816

Qy 2140 AAGATCTTAAACATCAATCAAAATAGAGATGCCAAATTCGAGTCAGTATTTTCTATCCC 2199

Db 8817 AAGATCTTAAACATCAATCAAAATAGAGATGCCAAATTCGAGTCAGTATTTTCTATCCC 8876

Qy 2200 AATACATTAACCGCTTAGTGAAAAAAGCTAAACAATATTTTGAATATAATAAATAATA 2259

Db 8877 AATACATTAACCGCTTAGTGAAAAAAGCTAAACAATATTTTGAATATAATAAATAATA 8936

Qy 2260 TTCGTTATTTATCTACATGTTGTAAGAATCATCTTACACAGACATCAAAAAGAAATA 2319

Db 8937 TTCGTTATTTATCTACATGTTGTAAGAATCATCTTACACAGACATCAAAAAGAAATA 8996

Qy 2320 TTGGCTTCTATCATAGCACCAGTGAATATTTTACTTAATATAGACATCTCATATTAC 2379

Db 8997 CTAGCCTTCTATCATAAACATCAAGTGAATATTTTACTTAAATAATGATATCTCATATTAC 9056

Qy 2380 ACGAGTAAATAGACTAATAAAGCTGAGGCACATTTAAGTAAATATTAATAAATTAAGTCAG 2439

Db 9057 ACGAGTAAATAGACTAATAAAGCTGAGGCACATTTAAGTAAATATTAATAAATTAAGTCAG 9116

Qy 2440 TTAATCTAAATGTTGAATACATCATTTTGTGAATCATGACAGCCTATTCGTTAAAAAT 2499

Db 9117 TTAATCTAAATGTTGAATACATCATTTTGTGAATCATGACAGCCTATTCGTTAAAAAT 9176

Qy 2500 GACAGCTATGCTTATAGAAAAATATGATGTCGGCATGAATTTCTCGACATTAACACAT 2559

Db 9177 GACAGCTATGCTTATAGAAAAATATGATGTCGGCATGAATTTCTCGACATTAACACAT 9236

Qy 2560 GATTGGATCGAGAAATCAATGCGCATCCACATTTTAAAAAGCTGATTTAAAAACCTATTTT 2619

Db 9237 GATTGGATCGAGAAATCAATGCGCATCCACATTTTAAAAAGCTGATTTAAAAACCTATTTT 9296

Qy 2620 AATGACATGACTTAAAGAGATGAAATGTGAAGGGGCATCACAGGATGTTTATGAAG 2679

Db 9297 AATGACATGACTTAAAGAGATGAAATGTGAAGGGGCATCACAGGATGTTTATGAAG 9356

Qy 2680 TATGCGCTACCGCATGAGCTTCTGACGATTATTAAAGAGTTCATCACATCTCCCAATCA 2739

Db 9357 TATGCGCTAGCGCATGAGCTTCTGACGATTATTAAAGAAAGTTCATCACATCTCCCAATCA 9416

Qy 2740 ATTGATAGTGTGCCAGAAATATAACACTGAGGATATTTGGTTCCAAATTTGCACCTTTTAATC 2799

Db 9417 ATTGATAGTGTGCCAGAAATATAACACTGAGGATATTTGGTTCCAAATTTGCACCTTTTAATC 9476

Qy 2800 TTAGAAAAAGAAACCGGCATGTTTATTAATAAACAATCGACCTGACTTATATGCGCTTGG 2859

Db 9477 TTAGAAAAAGAAACCGGCATGTTTATTAATAAACAATCGACCTGACTTATATGCGCTTGG 9536

Qy 2860 GAACCAAAATTTACAATGAGCAAAATGAACAAATTTCAAAGTGCAGAAAAAGCGCAAAATATC 2919

Db 9537 GAACCAAAATTTACAATGAGCAAAATGAACAAATTTGAAAGTGCAGAAAAAGAGAGAAAAATATA 9596

Qy 2920 CCGGTAAACAAGTTCATTATTAAATAGTAAACGCTATAAACAATTTGCAATTTTATTAATAA 2979

Db 9597 CCGGTAAACAAGTTCATTATTAAATAGTAAACGCTATAAACAATTTGCAATTTTATTAATAA 9656

RESULT 5

AF036004

LOCUS

DEFINITION

AF036004

VERSION

AF036004.1

GI:3043922

KEYWORDS

SOURCE

ORGANISM

Pasteurella multocida.

Pasteurella multocida.

Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

Pasteurella.

1 (bases 1 to 2937)

DeAngelis,P.L., Jing,W., Drake,R.R. and Achyuthan,A.M.

Identification and molecular cloning of a unique hyaluronan

synthase from Pasteurella multocida

J. Biol. Chem. 273 (14), 8454-8458 (1998)

98192845

PUBMED

9525958

2 (bases 1 to 2937)

DeAngelis,P.L., Jing,W. and Achyuthan,A.M.

Direct Submission

Submitted (26-NOV-1997) Biochem. & Molec. Biol., Univ. of Oklahoma

Health Sciences Center, 940 Stanton L. Young Blvd., Oklahoma City,

OK 73104, USA

Location/Qualifiers

1. .2937

/organism="Pasteurella multocida"

/strain="P-1059; ATCC 15742"

/db_xref="ATCC:15742"

/db_xref="taxon:747"

/note="Carter Type A strain"

19. .2937

/gene="PmHAS"

/note="gene identified by transposon insertional

mutagenesis and verified by functional expression in

heterologous cells"

19. .2937

/gene="PmHAS"

/function="polymerizes hyaluronan (HA, hyaluronate,

hyaluronic acid) polysaccharide using UDP-GlcA and

UDP-GlcNAc precursors"

/note="glycosyltransferase; HA synthase (synthetase);

membrane-bound enzyme"

/codon_start=1

/transl_table=11

/product="hyaluronan synthase"

/protein_id="AAC38318.1"

/db_xref="GI:3043923"

/translation="MNTLSQAIKAYNSNDYQALKLFKSAEYGRKIVEFOITKQOE

KLSAHPSVNSAHLVSYNKEEVNCDSPLDIATQLLLSNVKKVLSDSEKNTLKNKKL

LTEKSENAEYRAVALYPKDPKDLVLAPLDHVNDFTWYKKRKLRLGKPEQHVGV SIIYTFNRPAILSLITACLUNQKTHPIFFVIVDDSGOEDLSPIBOVENKLDIRVY ROKONGFOASARNMGRUKLADIFGLDDCDMAFNPLWVHSYIAELLEDDELLIIGPR KYIDTHIDPKDFLNNASLESLEPVTNNVAAGKEGTVSLDWRLGEQFEXTENLRLS DSRPFFAAGNAVAFKWLKNSGFFDEFNWGGEDVEFYRLFRYSFFFTIDIGIMA YHOBPPGKENTDREAGNIITLDIMREKVPYIYRKLPLIEDSHINRVPLSYIIPAYN CANVORCVDSALNCTVVDLEVCICNDGSTONTLEVINKLYGNPNRPRVIMSKPNPGIA SASNAVSAFAGYIIGOLDSDYLEPDAVELCLKEFLDKTLACVYITNRNVPDGSJ IANGYNWPEFSREKLTITAMIAHHFRMTIRAWHLTDGNEKLENAVDYDMELKLESEVG KFKHLKICYNRVLHGDNTSKKLGQKRNHFNVPVYNQNGQITTYNNDEDDLDDES RKYIFNKATYQEEEDILDKIKIQMKDAKIAVSIFYPNTLNGVLKKNLIEYNNKI FVILVHVBKNHLTPDKKEILAFKHQVNNILLNNDISYTSNRLLIKTEAHLNKNKL SOLNLCVSIIPDNHDSLVKNDVYMKYDVMNFSALTDHMTKEKNANHPFKKLI KTYFNNDLKSNNVKGASOGEMFTYALAEHLTIIEKVIITSCOSIDSXPENXTDWF QFALLILEKKYGHVFNKSTLTYMPWERKLOWTNEQIESAKRGENIPVYNKFIINSITL				
BASE COUNT	1077 a	482 c	509 g	869 t
ORIGIN				
Query Match	76.6%; Score 2283.4; DB 1; Length 2937;			
Best Local Similarity	86.5%; Pred. No. 0;			
Matches 2540; Conservative	0; Mismatches 376; Indels 21; Gaps 1;			
QY	43	ATTTTAAAGGAAGAAATCAATACATTATCAACAGCAATAAAGCATATAACGCAAT	102	
Db	1	ATTTTAAAGGACAGAAATGAATACATTATCAACAGCAATAAAGCATATAACGCAAT	60	
QY	103	GACTATGAATAGCACHCAAAATTTATTTGAGAAGTCCTGCTGAAACCTACGGCGCAAAAATC	162	
Db	61	GACTATCAATTAGCACHCAAAATTTATTTGAAAAGTCGGCGGAAATCTATGGACGGAAAAATT	120	
QY	163	GTTGAATCCAAATATCAAAATGTAAGAAAAAACTCTCGACCAATTC-----	209	
Db	121	GTTGAATTTCAAATATCAAAATGCCAGAAAAAACTCTCAGCACATCTCTGTTAAATTC	180	
QY	210	-----TTATGTAAGTGAAGATAAATAAAACAGTGTTCGCGATAGCTCATTAGATATC	261	
Db	181	GCACATCTTCTGTAAATAAAGAAAGAAAGTCAATGTTTCGCGATAGTCGCTAGATATT	240	
QY	262	GCAACAGCTCTTACTTCCAGGTAAAAAAATTAATCTATCCGATCAGAAAAAAAC	321	
Db	241	GCAACAACTGTTACTTTCACAGGTAAAAAAATTAGTACTTTCTGACTCGGAAAAAAAC	300	
QY	322	AGTTTAAAAATAATGGAATCTATCACTGGGAAAAAATCGGAGAACGCGAGAAATCAGA	381	
Db	301	ACGTTAAAAATAATGGAATTCCTCACTGAGAAGAAATCTGAAATCGCGAGGTAAGA	360	
QY	382	AAGGTGGAACTAGTACCAGAAATTTTCCCTAAAGATCTTGTCTTGTCTCCATTGCCAGAT	441	
Db	361	CGCGTCGCCCTTGTACCAAAAGATTTTCCCAAGATCTGGTTTTACGCGCTTTACCTGAT	420	
QY	442	CATGTTAATGATTTTACATGTCACAAAATCGAAAAAAAGCTTAGGTATAAAGCCTGTA	501	
Db	421	CATGTTAATGATTTTACATGTGACAAAAGCGAAAGAAAGACTTGGCATAAAACCTGNA	480	
QY	502	AATAAGAATATCGGTCTTCTTATTATTATTCCTACATTTAATCGTCGGCTATTTTAGAT	561	
Db	481	CATCAACATGTTGGTCTTCTATTATTCGTTTACAACATTTCAATCGACCAATTTTATCG	540	
QY	562	ATAAGTTAGCCTGTTTGGTCAATCAGAAAAACAACTACCCATTTGAAGTCGTTGTGCA	621	
Db	541	ATTACATTTAGCCTGTTTGTAGTAAACCAAAAAACACATTACCCGCTTTGAAGTTATCGTG	600	
QY	622	GATGATGCTAGTAAAGAAAACTTACTTACCATTTGTGCAAAAAATACGAAAAAACTTGAC	681	
Db	601	GATGATGCTAGTCAGGAAGATCTATCACCGATCATTCGCCAATATGAAAAATAAATTTGAT	660	
QY	682	ATAAAGTATGTAAAGACAAAAAGATTATGGATATCAATTTGTGTGAGTCAGAACTTAGGT	741	
Db	661	ATTGCGTACGTCAGACAAAAAGATACGGTTTTTCAAGCCAGTGTCCGCTCGGAATATGGGA	720	
QY	742	TTACGTCACGCAAGATGATTTTTGTCTCGATTCTTACCTGCGATATGGCACCACACAA	801	

Db	721	TTACGCTTAGCAAAATATACATTTATTTGGCTTACTCGACTGTGATATGGCGCAAAATCCA	780
QY	802	TTATGGTTTCATCTTATCTTACAGAACTATTAGAAGACAAATGATATGTTTAAATTTGA	861
Db	781	TTATGGTTTCATCTTATGTTGCAGAGCTATTAGAAGATGATGATTAAACAATCATTTGT	840
QY	862	CCTAGAAAAATATGTGGATCTACTAATAATATTACCGGAGAACAAATTCCTTAAACGATCC	921
Db	841	CCAAGAAAAATACATCGATACACAAATATTGACCCAAAAGACTTCTTAATAACGGGAGT	900
QY	922	TTAATAAGATCACTACTCTGAAACCGCTACAAAATAACAATCCTTCGATTATACATCAAAAG	981
Db	901	TTGCTTGAATCATTCACAGAGGTGAAACCAATAATAGTGTCCCGCAAAAGGGAAGGA	960
QY	982	AATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAAACCGAATATCTACGCTCTATGT	1041
Db	961	ACAGTTCTCTGGATTGGCGCTTTAGAACAAATTCGAAAAAACAGAAAAATCTCCGCTTATCC	1020
QY	1042	GATTCCTCGTTTCGTTTATTTTGGGGTAAATGTTTGCATTTTCTAAAGAAATGGCTAAAT	1101
Db	1021	GATTCGCTTTTCGTTTTCGCGGGTAAATGTTTGCCTTAAATAATGGCTAAAT	1080
QY	1102	AAAGTAGTTGGTTTCGATGAAGAATTTAATCATTTGGGGGGCGGAAGATGTAGAAATTTGGT	1161
Db	1081	AAATCCGGTTTCTTGTAGAGAGAAATTTAATCATCTGCGGTGAGAGAGATGTGGAATTTGA	1140
QY	1162	TACAGATTATTGCAAAAGGCTGTTTTTTCAGAGTAAATGACGGCGGAATGGCCATCAT	1221
Db	1141	TATCGCTTATTCGTTTACGCTAGTTTCTTTAAAACTATTGATGGCATATTATGGCTTACCA	1200
QY	1222	CAAGAACCCCTGGTAAAGAAATGAAACAGACGCGAAGCTGCTAAAAGTATTACGCTT	1281
Db	1201	CAAGACGACGAGTAAAGAAATGAAACCGGATCGTGAAGCGGGAATAATATTACGGCTC	1260
QY	1282	AAAATCTGAAAGAAAAAGTACCTTACATCTATAGAAGCTTTTACCATAAAGAAATTC	1341
Db	1261	GATATTATGACAGAAAAAGTCCCTTATCTATAGAAAACCTTTACCAATAAGAGATTCG	1320
QY	1342	CATATTCATAGAATACCTTTTATTTTATATCCCGCTTATACTGTGCAAAATTTAT	1401
Db	1321	CATATCAATAGAGTACCTTTTATTTTATATCCAGCTTATACTGTGCAAACTAT	1380
QY	1402	ATTCAAGATGCTGTAGATAGTGTCTTAAATCAAACTGTTGTCGATCTCGAGGTTTGTATT	1461
Db	1381	ATTCAAGCTTCCGTAGATAGTGTGCTGACTGACTGACTGCTGTTGTTGATCTCGAGGTTT	1440
QY	1462	TGTAACGATGTTTCAACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGGTAAATAT	1521
Db	1441	TGTAACGATGTTTCAACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGGTAAATAT	1500
QY	1522	CCTAGGGTACGATCATGCTTAAACCAATGGCGGAATAGCTCAGCATCAAAATCGAGCC	1581
Db	1501	CCTAGGGTACGATCATGCTTAAACCAATGGCGGAATAGCTCAGCATCAAAATCGAGCC	1560
QY	1582	GTTCTTTTGTAAAGTTTATACATTTGGGAGTTAGATTTCAGATGATTATCTTGAGCCT	1641
Db	1561	GTTCTTTTGTAAAGTTTATACATTTGGGAGTTAGATTTCAGATGATTATCTTGAGCCT	1620
QY	1642	GATGAGTTGAACCTGTTTAAAAGAAATTTTAAAAGATAAAACCGTAGCTTGTGTTAT	1701
Db	1621	GATGAGTTGAACCTGTTTAAAAGAAATTTTAAAAGATAAAACCGTAGCTTGTGTTAT	1680
QY	1702	ACCCTAATAGAAACGTCATCCGGATGTTAGCTTAAATGCTTAATGGTTTACAAATTTGGCCA	1761
Db	1681	ACCCTAATAGAAACGTCATCCGGATGTTAGCTTAAATGCTTAATGGTTTACAAATTTGGCCA	1740
QY	1762	GAATTTTCCAGGAAAAACTCACACGGCTATGATTGCTCACCAATTTTGAATGTTTACG	1821
Db	1741	GAATTTTCCAGGAAAAACTCACACGGCTATGATTGCTCACCAATTTTGAATGTTTACG	1800
QY	1822	ATTAGAGCTTGGCATTTTAAACGGATGATTAAACGAAATATTGCAAAACCGCTGGATTA	1881
Db	1801	ATTAGAGCTTGGCATTTTAACTGATGATTCATGAAATAATTTGAAATGCTGCTGATGAT	1860

RESULT 6	AF237926	2919 bp	DNA	linear	BCT 08-MAY-2000
LOCUS	Pasteurella multocida hyaluronan synthase (has) gene, complete cds.				
DEFINITION	Pasteurella multocida hyaluronan synthase (has) gene, complete cds.				
ACCESSION	AF237926				
VERSION	AF237926.1	GI:7716512			
KEYWORDS					
SOURCE	Pasteurella multocida.				
ORGANISM	Pasteurella multocida				
REFERENCE	1 (bases 1 to 2919)				
AUTHORS	Fuller,T.E., Kennedy,M.J. and Lowery,D.E.				
TITLE	Identification of Pasteurella multocida virulence genes in a				
JOURNAL	septicemic mouse model using signature-tagged mutagenesis				
AUTHORS	Unpublished				
TITLE	2 (bases 1 to 2919)				
JOURNAL	Fuller,T.E., Kennedy,M.J. and Lowery,D.E.				
AUTHORS	Direct Submission				
TITLE	Submitted (24-FEB-2000) Discovery Research, Pharmacia & Upjohn				
JOURNAL	Animal Health, 7923-25-434, 7000 Portage Road, Kalamazoo, MI				
FEATURES	49001-0199, USA				
source	Location/Qualifiers				
gene	1. .2919				
CDS	2. .2919				
	3. .2919				
	4. .2919				
	5. .2919				
	6. .2919				
	7. .2919				
	8. .2919				
	9. .2919				
	10. .2919				
	11. .2919				
	12. .2919				
	13. .2919				
	14. .2919				
	15. .2919				
	16. .2919				
	17. .2919				
	18. .2919				
	19. .2919				
	20. .2919				
	21. .2919				
	22. .2919				
	23. .2919				
	24. .2919				
	25. .2919				
	26. .2919				
	27. .2919				
	28. .2919				
	29. .2919				
	30. .2919				
	31. .2919				
	32. .2919				
	33. .2919				
	34. .2919				
	35. .2919				
	36. .2919				
	37. .2919				
	38. .2919				
	39. .2919				
	40. .2919				
	41. .2919				
	42. .2919				
	43. .2919				
	44. .2919				
	45. .2919				
	46. .2919				
	47. .2919				
	48. .2919				
	49. .2919				
	50. .2919				
	51. .2919				
	52. .2919				
	53. .2919				
	54. .2919				
	55. .2919				
	56. .2919				
	57. .2919				
	58. .2919				
	59. .2919				
	60. .2919				
	61. .2919				
	62. .2919				
	63. .2919				
	64. .2919				
	65. .2919				
	66. .2919				
	67. .2919				
	68. .2919				
	69. .2919				
	70. .2919				
	71. .2919				
	72. .2919				
	73. .2919				
	74. .2919				
	75. .2919				
	76. .2919				
	77. .2919				
	78. .2919				
	79. .2919				
	80. .2919				
	81. .2919				

Qy 220 GAAGATAAAAAACAGTGTGATAGCTCATTAGATATCGCAACACAGCTCTTACTT 279
Dy 181 AAAGAGAAAAGTCAATGTTGGATAGTCGGTTAGATATGCAACACACTGTTACTT 240
Qy 280 TCCACGCTAAAAAATTAACCTTATCCGAATCAGAAAAACAGTTTAAAAAATAAATGG 339
Dy 241 TCCACGCTAAAAAATTAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 340 AAATCTACTGCGGAAAAAATCGGAGAACGAGAACGAGAAATCAGAAAGGTGGAACCTAGTACCC 399
Dy 301 AAATGCTCACTGAGAGAAATCTGAAATCGGAGGTGAAGGCGGTGCGCCCTGTAGCCA 360
Qy 400 AAAGATTTTCTTAAGATCTGTTCTGCTCCATTGCCAGATCATGTTAAATGATTTTACA 459
Dy 361 AAAGATTTTCCCAAGATCTGTTTACGCTTTACCTGATCATGTTAAATGATTTTACA 420
Qy 460 TGGTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCTGTAATTAAGAAATACGCTCTT 519
Dy 421 TGGTACAAAAAGCGAAAAAGACTTGGCATAAAACTGAACATCAACATGTTGGTCTT 480
Qy 520 TCTATTATATTCCTACATTTAATCGTAGCCGTAATTTTAGATATATACGTTAGCCCTGTTG 579
Dy 481 TCTATTATCGTTACACATTTCAATCGACGACCAATTTTATCGATTACATTTAGCCCTGTTT 540
Qy 580 GTCAATCAGAAAAAATACCCATTTCGAAGTCGTTGTCAGATGATGGTAGTAAGGAA 639
Dy 541 GTAACCAAAAAACATTTACCCGTTTGAAGTTATCGTGACAGATGATGGTAGTCAGAA 600
Qy 640 AACTTACTTACCATTTGCGAAAAATACGAAACAAAACTTGACATAAAGATATGTAAGACAA 699
Dy 601 GATCTTATCACCGATCATTCGCGCAATTAAGAAATTAATTCGATTCGCTAGTCAGACAA 660
Qy 700 AAAGATTTATGATATCAATGTCGCTCAGAACTTAGTTTACGTACAGCAAGATAT 759
Dy 661 AAAGATTAACGGTTTCAAGCAGTCCGCTCGAATATGGGATTTAGCTTTAGCAAAATAT 720
Qy 760 GATTTTGTCTCGATTCTAGACTGCGATATGGCACCACAACTTAATGTTGTTCACTTTAT 819
Dy 721 GACTTTATTTGGCTTACTCGACTGATATGGCGCAAAATCCATTATGGTTCACTTCTAT 780
Qy 820 CTTACAGAACTATTAGAGACAACTGATTTGTTTAAATGGACCTAGAAAAATATGTGGAT 879
Dy 781 GTTGCAAGACTATTAGAGATGATGATTAACAATCATTTGTTCCAGAAAAATACATCAT 840
Qy 880 ACTCATATATTACCGCAGAACATTCCTTAACGATCCATATTTAATAGAATCACTACT 939
Dy 841 ACACAACTATTGACCCAAAGACTTCTTAATAACGGGAGTTTCTGTTGAATCATTTACA 900
Qy 940 GAAACCGCTACAAATAACAACTCTTCGATTACATCAAAAGGAAATATATCGTTGGATTGG 999
Dy 901 GAAGTGAAACCAATAATAGTGTCCGCAAAAGGGAGGAACAGTTTCTCTGGATTGG 960
Qy 1000 AGATTAGAACATTTCAAAAAACCGATTAATCTACGCTCTATGATATTCGCTTTCGTTAT 1059
Dy 961 CGCTTAGAACATTTGAAAAACAGAAAAATCTCCGCTTATCGGATTCGCTTTCGCTTTT 1020
Qy 1060 TTTGTTGCGGTAATGTTGCTATTTTAAAGATCGTAAATAAAGTAGTTGTTTCGAT 1119
Dy 1021 TTTGCGGCGGTAATGTTGCTTTCGCTTAAATAAGCTTAAATAAATCGGTTTCTTTGAT 1080
Qy 1120 GAAGAAATTTAATCATTTGGGGGCGGAAAGATGTAGAAATTTGGTTTACAGATATTTGCCAAA 1179
Dy 1081 GAGGAATTTAATCACTGGGGTGAAGAGATGTGGAATTTGGATATCGCTTATTCGGTTTAC 1140
Qy 1180 GCGTGTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCATCAAGAACCCCTGGTAAA 1239
Dy 1141 GGTAGTGTCTTTTAAAACTATTGATGGCAATTTAGGCTTACCTACCAATCAAGAGCCACGAGTAAA 1200
Qy 1240 GAAATGAACAGAACGCGAAGCTGTAAAAGTATTACGCTTAAATTTGTGAAAGAAAG 1299
Dy 1201 GAAATGAACCCGATCGTGAACGGGAAAAAATATTTACGCTCGATATTATGAGAGAAAG 1260

Qy 1300 GTACCTTACATCTATAGAAAGCTTTTCCAAATAGAAAGATTCACATATTTCATATAGATACCT 1359
Dy 1261 GTCCCTTATATCTATAGAAAACTTTTCCAAATAGAAAGATTCGCATATCAATAGAGTACCT 1320
Qy 1360 TTAGTTTCTATTTATATCCCGCTTATAACCTGCAAAATTAATATCAAGATGTTAGAT 1419
Dy 1321 TTAGTTTCAATTTATATCCAGCTTATACTGTGCAAACTATATTCAAGGTTGCGGTAGAT 1380
Qy 1420 AGTGTCTTTAATCAAACTGTTGTCGATCTCGAGGTTGTTGTTGTTAAACGATGTTCAACA 1479
Dy 1381 AGTGCACTGAATCAGACCTGTTGTTGATCTCGAGGTTGTTGTTGTTGAACGATGTTCAACA 1440
Qy 1480 GATAATACCTTAGAAGTGAATCAATAGCTTTATGTTAATAATCTTAGGTTAGGTTAGCATCATG 1539
Dy 1441 GATAATACCTTAGAAGTGAATCAATAGCTTTATGTTAATAATCTTAGGTTAGGTTAGCATCATG 1500
Qy 1540 TCTAAACCAATGGCGGAATAGCCTCAGATCAAAATGACGCGTTTCTTTTCTTAAAGT 1599
Dy 1501 TCTAAACCAATGGCGGAATAGCCTCAGATCAAAATGACGCGTTTCTTTTCTTAAAGT 1560
Qy 1600 TATTACATTTGGCAGCTTAGATTTCAGATGATTTATCTTGAGCCTGATGCGATTTCAACTGTCT 1659
Dy 1561 TATTACATTTGGCAGCTTAGATTTCAGATGATTTATCTTGAGCCTGATGCGATTTCAACTGTCT 1620
Qy 1660 TTAAAGAAATTTTAAAGATAAAACGCTAGCTGCTGTTTATACCACTAATAGAAACGTC 1719
Dy 1621 TTAAAGAAATTTTAAAGATAAAACGCTAGCTGCTGTTTATACCACTAATAGAAACGTC 1680
Qy 1720 AATCCGGATGGTACCTTAATCGCTAATGGTTACAAATGGCCAGAAATTTTCACGAGAAAAA 1779
Dy 1681 AATCCGGATGGTACCTTAATCGCTAATGGTTACAAATGGCCAGAAATTTTCACGAGAAAAA 1740
Qy 1780 CTCACACGCGCTATGATTGCTCACCATTTTAGAAATGTTTACGATTTAGAGCTTTGGCATTTA 1839
Dy 1741 CTCACACGCGCTATGATTGCTCACCATTTTAGAAATGTTTACGATTTAGAGCTTTGGCATTTA 1800
Qy 1840 ACGGATGATTTTAAAGAAATTAAGAAACGCGGTGGATTTAGACATTTCTCTTAAACTC 1899
Dy 1801 ACTGATGATTTCAATGAAAAATTTGAAATGCGGTAGACTATGACATTTCTCTCAAACTC 1860
Qy 1900 AGTGAAGTTGGAATTTTAAACATCTTAATAAATCTGCTATAACCGCTATTACATGCT 1959
Dy 1861 AGTGAAGTTGGAATTTTAAACATCTTAATAAATCTGCTATAACCGCTATTACATGCT 1920
Qy 1960 GATAACACATCATTAAGAACTCGSCATTTCAAAAGAAAAACCATTTTGTGTAGTCAAT 2019
Dy 1921 GATAACACATCAATTAAGAACTTGSCATTTCAAAAGAAAAACCATTTTGTGTAGTCAAT 1980
Qy 2020 CAGTCATTAATAGACAGGCATCAATTAATTAATATGACAAATTTGATGATTTAGAT 2079
Dy 1981 CAGTCATTAATAGACAGGCATCAATTAATTAATATGACAAATTTGATGATTTAGAT 2040
Qy 2080 GAAAGTAGAACTATATCTTCAATAAACCGCTGAATATCAAGAGAAATGATATGTTA 2139
Dy 2041 GAAAGTAGAACTATATCTTCAATAAACCGCTGAATATCAAGAGAGATGATATGTTA 2100
Qy 2140 AAAGATCTTAAACTTATCAAAATTAAGATGCAAAATCGCACTAGTATTTTCTATCCC 2199
Dy 2101 AAAGATATTAATCATCCAGATTAAGATGCAAAATCGCACTAGTATTTTCTATCCC 2160
Qy 2200 AATACATTAACGGCTTAGTGAATAAACTTAACAATATTATTGAATATTAATAAATAATA 2259
Dy 2161 AATACATTAACGGCTTAGTGAATAAACTTAACAATATTATTGAATATTAATAAATAATA 2220
Qy 2260 TTTGCTTATTATCTCATGTTGATAGAACTATCTTACACACAGACATCAAAAGAAAAATA 2319
Dy 2221 TTTGCTTATTGTTCTACATGTTGATAGAACTATCTTACACAGATATCAAAAGAAAAATA 2280
Qy 2320 TTTGCTTATTATCATAAGCACCAGTGAATTTTACTAAATAATGACATCTCATATTAC 2379
Dy 2281 CTAGCCTTCTCATATAAACATCAAGTGAATTTTACTAAATAATCATATCTCATATTAC 2340
Qy 2380 ACGAGTAATAGACTTAATAAACTGAGGCACATTTTAAGTAAATTAATAAATAAGTCAG 2439

```
|||||
Db 2341 ACAGTAAATAGATTAATAAAACCTGAGCGCATTTAAAGTAATAATAAAATTAAGTCAG 2400
Qy 2440 TTAATCTAAATCTGAATACATCATTTTCATATAATCATGACGCCCTATTCGTTAAAAAT 2499
Db 2401 TTAATCTAAATCTGAATACATCATTTTCATATAATCATGACGCCCTATTCGTTAAAAAT 2460
Qy 2500 GACAGCTATGCTTATATCAAAAAATATGATCGCGCATGAATTTCTCAGCATTAACACAT 2559
Db 2461 GACAGCTATGCTTATATCAAAAAATATGATCGCGCATGAATTTCTCAGCATTAACACAT 2520
Qy 2560 GATTGGATCGAGAAAAATCAATGCGCATCCACCATTTTAAAAAGCTGATTTAAACCTATTTT 2619
Db 2521 GATTGGATCGAGAAAAATCAATGCGCATCCACCATTTTAAAAAGCTGATTTAAACCTATTTT 2580
Qy 2620 AATGACAAATGACTTAAAGATGATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAG 2679
Db 2581 AATGACAAATGACTTAAAGATGATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAG 2640
Qy 2680 TATCGCTACCGCATGAGCTTCTGACGATTTATTAAGAAGTCATCACATCTCGCAATCA 2739
Db 2641 TATCGCTACCGCATGAGCTTCTGACGATTTATTAAGAAGTCATCACATCTGCGAGTCA 2700
Qy 2740 ATTGATAGTGTGCCAGAAATATAACATGAGGATATTTGGTTCCAAATTTGCATTTTAATC 2799
Db 2701 ATTGATAGTGTGCCAGAAATATAACATGAGGATATTTGGTTCCAAATTTGCATTTTAATC 2760
Qy 2800 TTACAAAGAAACCGGCCATGTTATTAATAAATCATCGACCCCTGACTTATATGCGCTGG 2859
Db 2761 TTACAAAGAAACCGGCCATGTTATTAATAAATCATCGACCCCTGACTTATATGCGCTGG 2820
Qy 2860 GAACGAAATTTACAAATGACAAATGAACAAATTTCAAGTGCAGAAAAAGCGAAATATC 2919
Db 2821 GAACGAAATTTACAAATGACAAATGAACAAATTTCAAGTGCAGAAAAAGCGAAATATC 2880
Qy 2920 CCCGTTAACAAAGTTCATTATTAATAGTATTAACGCTATAA 2958
Db 2881 CCTGTTAACAAAGTTCATTATTAATAGTATTAACGCTATAA 2919

RESULT 7
AB079602/c
LOCUS
DEFINITION
Escherichia coli kpsS, kfoG, kfoF, kfoB, kfoD, kfoI, kfoJ, kfoC, kfoB,
kfoA, kpsT genes for KpsS hypothetical protein, predicted
glycosyltransferase, UDP-glucose dehydrogenase, KfoE hypothetical
protein, KfoB hypothetical protein, IS hypothetical protein,
chondroitin polymerase, KfoB hypothetical protein, UDP-glucose
4-epimerase, ABC transporter, partial and complete cds.
ACCESSION
AB079602
VERSION
AB079602.1 GI:21326777
KEYWORDS
Escherichia coli (strain:K4) DNA.
ORGANISM
Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1
Ninomiya,T., Sugliura,N., Tawada,A., Sugimoto,K., Watanabe,H. and
Kimata,K.
Molecular Cloning and Characterization of Chondroitin Polymerase
from Escherichia coli strain K4
J. Biol. Chem. 277 (24), 21567-21575 (2002)
11943778
PUBMED
REFERENCE
2
(bases 1 to 14483)
Sugliura,N., Kimata,K., Ninomiya,T. and Watanabe,H.
Direct Submission
Submitted (07-FEB-2002) Nobuo Sugliura, Aichi Medical University,
Institute for Molecular Science of Medicine; 21, Yazako, Nagakute,
Aichi 480-1195, Japan (E-mail:nsugliura@aichi-med-u.ac.jp,
URL:http://www2.aichi-med-u.ac.jp/ims/index-e.html,
Tel:81-52-284-4811(ex.2087), Fax:81-561-63-3532)
Location/Qualifiers
1. .14483
source
```

```
/organism="Escherichia coli"
/strain="K4"
/serotype="O5:K4(L):H4"
/db_xref="taxon:562"
1. .233
/gene="kpsS"
<1. .233
/gene="kpsS"
/codon_start=3
/transl_table=11
/product="kpsS hypothetical protein"
/protein_id="BAC00517.1"
/db_xref="GI:21326778"
/translation="ISALHNKPLKVMGNALYDIKGLTYQCHLQHFQWQADFKNKPLF
KKPRGYLLVKTQVNAVYVYGEIMFNKIKKKR"
complement(393. .1430)
/gene="kfoG"
complement(393. .1430)
/gene="kfoG"
/codon_start=1
/transl_table=11
/product="predicted glycosyltransferase"
/protein_id="BAC00518.1"
/db_xref="GI:21326779"
/translation="MPVILYIPLIYVYCKNNAFRETAISLNRNLEYKYPGLSPVAL
TNSSPSPCNIIITSIPYAOPTPLISILMTTYNGRRVENAVISLLNOTYRSPFLII
VDDASTDTLFLRLALAKDPRIKIIISLPQNVGYAAKRIGLQAKGFVTHDSDDW
SHPKLFRQISPLLNPKLIGSISDWVRLQDNGIFYARAVYPLKRLNPSLLFRADY
LQAGVWDCVKTGADSEFTAKLIFGDSITVHRKILPTLGLSHRTDSLWNSPTTITS
QGSTPDRQYWSWSRWHQIALRNKESLYIGNSDFTNKNRPFSPDSILVDTNAIKTA
LQSAHVNFSTI"
complement(1862. .3031)
/gene="kfoF"
complement(1862. .3031)
/gene="kfoF"
/codon_start=1
/transl_table=11
/product="UDP-glucose dehydrogenase"
/protein_id="BAC00519.1"
/db_xref="GI:21326780"
/translation="MKIAGVGVYGISIAILLSQKHDIIALDIDPKKVLINKKISP
ICDPEIQKLSNRKLNLYATTEKYEAYRDADYVIAITPTNDPINNFDTLUSVSAC
DVLISFNPAIIIKSTVPVGTGERLKDNTNNIIFSPFREGKALYDNLVPSRIVY
GESSERARKFAELSEGAIKKIDIPILLTDSPEAIAKLFANTYLAIRYAYNELDTYA
SVHGLDTQIIIEGVSLDPRIGQHVNPSEGVGCLPKDQKSLANYRDVQNLIQAI
VDANTRKDFVAEDILSRKPKVGVYIRLIMKAGSDNFPASSIOGVNKLKAGIIVV
YEVLPKEPYFGSYVERDINSFKERDVIVANRRITSELEDVSEKVTYTRDLFGVDS"
complement(3850. .5418)
/gene="kfoB"
complement(3850. .5418)
/gene="kfoB"
/codon_start=1
/transl_table=11
/product="KfoE hypothetical protein"
/protein_id="BAC00520.1"
/db_xref="GI:21326781"
/translation="MLLIIMSGSYVQDELGAFFGSTPPSPFLANKRLFKHVSILGHGD
HAIYVLPEDFVFDKHYEILLRNKVTMIPVDSNLTQGAIVTAWNLIGDKDDKGLQ
LFGDLTKKIPAGELVAKSPDENQWAIETELRAVSREDNNKVICGFSPFRKP
FFIRELYTSKFDAAALKKYHDSYSLASIVYSDWLDGFIINTYYSKSYQYTORAFNE
LCITTKSVIKSSSNEAKTEAKSKWFTETIPGELKIYTPMLLEFDHIRKSYKLEYLNT
TNELFVSRPLNNILTNILISCLDFIDLCREYHSIDFKNIDLODLFIERYVSKY
ITDLIDNPNAKNNNNISVINDILYDTNKFIPSELOYKTIHMGDLCFNSIENFRT
GRIQVDFPRLNHSGETSIYGDYDIKLSHSLUGLYDWLIAGYIINKNKNKTHSIE
FKINIDNKLFEIQSTFVSIIRKYSISEKSIYAMQIHLFLSNLPLHSDKKRQDNLFA
NAFLRYEIFKEAAV"
complement(5808. .7274)
/gene="kfoD"
complement(5808. .7274)
/gene="kfoD"
/codon_start=1
/transl_table=11
/product="KfoD hypothetical protein"
```

		repeat_region			
		gene			
		CDS			

Db	8973	TGCAATCATATGTCATCATTTTCAGGATGTTCCACAGCAAGACGATGGACACCTAACTGAAGG	8914
Qy	1848	ATTTTAACGAAAAATATTGAAAACGCCGTGGATTTATGACATGTTCCTTTAAACTCAGTGAAGT	1907
Db	8913	TTTCAACCAATCGATCATCAGCAACGCAGTTGATTACGATATGTATTTTAAAACTTTAGTGAAGT	8854
Qy	1908	TGGAAAAATTTTAAACATCTTTAATAAATCTGCTATTAACCGCGTATTACATGCGTGATAACAC	1967
Db	8853	TGGACCGTTCAAGCATATAAACAATAATTTGTTATATTCGCGTATTTGCGATGGTGAAAAATAC	8794
Qy	1968	ATCCATTTAAGAAACTCGCGCATTTCAAAAGAAAAACCATTTTCTGCTAGTCAATCAGTCAATT	2027
Db	8793	GTCTATATAAAGTTGGATATTCAAAAGGAATAATCTTTTAAAGTTGTTAACGAATCAATT	8734
Qy	2028	AAATAGACAAGGCATCAATTTATTAATTAATGACAAAATTTTGATGATTTAGATGAAGTAG	2087
Db	8733	AAGTAGCTAGGCATATAAAAAATATAAATATTACCACTTAACATACTTGAATGAATGTAG	8674
Qy	2088	AAAGTATATCTTCAATTAACACCGCTGAATATCAAGAGAAATG	2130
Db	8673	AAAATATACCTGGGAAAAAATAGAGAAATGATTTATAATATTATG	8631
RESULT 8			
AC007926/c			
LOCUS			
DEFINITION			
Trypanosoma brucei chromosome II clone RPC193-3H15, *** SEQUENCING			
IN PROGRESS ***, 1 ordered piece.			
ACCESSION			
AC007926.8 GI:14787210			
VERSION			
HTG; HTGS_PHASE2.			
KEYWORDS			
Trypanosoma brucei.			
SOURCE			
ORGANISM			
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;			
Trypanosoma.			
REFERENCE			
1 (bases 1 to 155204)			
El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,			
Peterson,L., Hou,L., Zhao,H., Mason,T., Millscher,J., Pai,G., Van			
Aken,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,			
Melville,S., White,O., Adams M.D., Donelson,J.E. and Fraser,C.M.			
Trypanosoma brucei GUTat10.1 RPC193-3H15 BAC genomic sequence			
Unpublished			
REFERENCE			
2 (bases 1 to 155204)			
El-Sayed,N.M., Khalak,H. and Adams,M.D.			
AUTHORS			
Direct Submission			
TITLE			
Submitted (28-JUN-1999) The Institute for Genomic Research, 9712			
JOURNAL			
Medical Center Dr, Rockville, MD 20850, USA			
COMMENT			
On Jul 17, 2001 this sequence version replaced gi:12746529.			
* NOTE: This is a 'working draft' sequence. It currently			
* consists of 1 contigs. Gaps between the contigs			
* are represented as runs of N. The order of the pieces			
* is believed to be correct as given, however the sizes			
* of the gaps between them are based on estimates that have			
* provided by the submittor.			
* This sequence will be replaced			
* by the finished sequence as soon as it is available and			
* the accession number will be preserved.			
* 1 155204: contig of 155204 bp in length.			
FEATURES			
source			
1. .155204			
/organism="Trypanosoma brucei"			
/isolate="GUTat10.1"			
/db_xref="taxon:5691"			
/chromosome="II"			
/clone="RPC193-3H15"			
BASE COUNT 48747 a 31603 c 32329 g 42525 t			
ORIGIN			
Query Match 3.3% Score 97.6; DB 2; Length 155204;			
Best Local Similarity 45.5%; Pred.No. 6.8e-07;			
Matches 346; Conservative 0; Mismatches 414; Indels 0; Gaps 0;			
Qy	1912	AAATTTAAACATCTTAATAAAATCTGCTATTAACCGCGTATTACATGTTGATTAACACATCC	1971

the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF
 STIALLYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
 J. Biol. Chem. 277 (1), 327-337 (2002)
 11689567
 2 (bases 1 to 11442)
 Gilbert, M., Michniewicz, J., Karwaski, M.-F., Cunningham, A. and
 Wakarchuk, W. W.
 Direct Submission
 Submitted (15-JUL-2001) Institute for Biological Sciences, National
 Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
 OR6, Canada

FEATURES

Location/Qualifiers

1..11442
 /organism="Campylobacter jejuni"
 /strain="ATCC 43438"
 /db_xref="ATCC:43438"
 /db_xref="taxon:197"
 /note="Penner-type strain O:10"

gene

CDS

<1..348
 /gene="aac"

<1..348
 /gene="aac"

/codon_start=1
 /transl_table=11
 /product="heptosyltransferase I"
 /protein_id="AAK91719.1"
 /db_xref="GI:15193215"

/translation="FAKEVNLNSGIDETHIELAPFSLBELMAFTKMDLIIGNSGP
 THLAPLNKASITIFGATPSYRNAFQTHINKIIDAGKIQNAKHIDKSDFCITRIEE
 DIFKLAKGLLNK"

338..1225
 /gene="htrB"

338..1225
 /gene="htrB"

/codon_start=1
 /transl_table=11
 /product="lipid A biosynthesis acyltransferase"
 /protein_id="AAK91720.1"
 /db_xref="GI:15193216"

/translation="MKNDRILSYLYILKFVTPMPECILHFLALIVARITFLNKK
 HRKIIDNLQICFPQYQOKERDKLSKIYENFAQGDICLQNTTKERILKNVFIN
 ENLLDALAKRPIIFTTAHYGNWEILSLAYAAKGAISIVGKKLKSELMYELSONR
 TQDIELIDKGGIKMSALKRAGLITLQDCVEKESVRLKFNKEVNYOMGASL
 IAQRNALIIPYAVKEDGFCIERFKVKDSQNASLEELTYQAQSCDEMIKKRPWEY
 FEHRFPASYNKEIYKDAK"

1222..2769
 /note="orf3"

/codon_start=1
 /transl_table=11
 /product="putative two-domain glycosyltransferase"
 /protein_id="AAK91721.1"
 /db_xref="GI:15193217"

/translation="MNLKQISALIIIVKNAQVLLLECLNSLKDFDEIILLNNESSDNTL
 NTANEFKDFANLHIYHSAPFIGALKNLALSAYKNDWILSDADAEVLENECKELKAN
 LEIQEDNIILALRGKNIKGWIKACGWPDYVVRINFTNFDNLVHESLVLPNSA
 KKIYKNGRHLAFRDIHSLIDKMOYISSLAKQNHKSGVLKANLRAFWTFFRNYF
 SVNGLFYGVKGTISYCNALGTFYKMKLYELQOKPKTCALITTYNOKERKLKLVLD
 KNGLAFLNLEVIADGSKEDTARLIEYQKDFPCPLKHIOEDGEGFRAAKSRNNAI
 KASKSYIILIDGMILEKDFVKNHLEFAKRVILQSGSVILNKSEETLKNNYSL
 AFNKGFKWORSIFLAKCVKFSKLFKFFKRSQLVKSGSKTCNNSPKSDFEAIEGFN
 EPIFGWGREDESEVARFLNFDGVSFKRLKFNALAHYIHEENSNMLEINHQIYLETIK
 NEKTTWR"

2773..3957
 /note="orf4"

/codon_start=1
 /transl_table=11
 /product="putative glycosyltransferase"
 /protein_id="AAK91722.1"
 /db_xref="GI:15193218"

/translation="MKKGVVPIYINVEYKVLKCELDVSNQTYTNLEIILVNDGSTDE
 NSNLAKVTLKDKRTLFDKKNGLSSARNVGLIEVFSGEYKLKNTQHTIKENSLIEF
 QLDGNPNYIKAYKSSQAFNNEKDLNTYPSIDYIIFLSDSNYKMLKCIIECVIRM
 KNVDVLFWDHDCITIEDNINKNKKHRTMEIFDFKKECIITPEKAYANRALSVGSRDISG

gene

CDS

CDS

CDS

WNGVIDNFLKQIKLKFINFIEDIHFGIILFASANKIYVLSQLYLCRLRANSISN
 HDKKTITRANSEYFKDIYETEGENAKENYILKAASRVITALKLFEFKDQKSENALA
 IKFTFLPCYAKKALMIKKFKDPLNLKQLVLKPFQTKLPYDIWKEWQIKKYLLII
 KI"
 4013..5071
 /gene="cgtA"
 4013..5071
 /gene="cgtA"
 /codon_start=1
 /transl_table=11
 /product="beta-1,4-N-acetylglactosaminyltransferase"
 /protein_id="AAK91723.1"
 /db_xref="GI:15193219"
 /translation="MLFQSYFVKIILFIPFKIRHKIKKTKLLKNQIORDKIDSYLPK
 KTLQIKNYNNEDLJLKNKAIIGGKHGKGFYNDKSKDPKSPINPWFARVKNKAEATL
 KASLESILPAIQGVIGVNDCTDSEIILEFCQYPSPIKPIYECIOIPKSEEN
 KLYSYNTVAFSPKDEWLKIDVHYDAKKLYKSFYIPRKNYHVISYRDIFFINE
 EKFYVRNKEGELLKAPGDLAIQNTNLFWKELLIEDDFKNTAKNNLENAKSYEIL
 KVRNRIYFTTELNNYHFFPKYRNKNDYKQLNWNVLDDDFIKNKIKELKQIDFKMLEY
 KTLKVKYKLTSSASDKI"
 complement(5043..5936)
 /gene="cgtB"
 complement(5043..5936)
 /gene="cgtB"
 /codon_start=1
 /transl_table=11
 /product="beta-1,3-galactosyltransferase"
 /protein_id="AAK91724.1"
 /db_xref="GI:15193220"
 /translation="WFKISIIPTYNVEQYIARATESCINOTFKNTEIIVDDGSDK
 SIDIVKEYAKDDRIKIHNEENKLRARYEGVKVANSFYINFLPDDYLELNACEE
 CMKILKNEIDILFFNAFVLENNKNIERKLNQKCYVRKDFLKLKLNPLFTVWA
 KVIKKELYLKAVGLISLENAKINMAEDVLLYPLINISNTIFHLSKLNLYQNNFSI
 TKTLTQNIKTNIQODNVLYLLKKQYVNFNLTLKLIYEFLLEIKYSLSSKRNVL
 CPKINIFPKIQKFYRLKLM"
 6025..6900
 /gene="cstII"
 6025..6900
 /gene="cstII"
 /codon_start=1
 /transl_table=11
 /product="alpha-2,3-/alpha-2,8-sialyltransferase"
 /protein_id="AAK91725.1"
 /db_xref="GI:15193221"
 /translation="MKKVIISGNPSLKEIDYSRLENDEDFVRCNOFYFEDKYLGGK
 FRAYFYNGLFFEQYTLKHLIQONQYETELIMCSNYNQAHLENENFVTFYDFPDA
 HLYGDFKQKLEFNAFKHEIYINQRTISGVYMCVAIATGAKLEYLSGIDFYQNS
 SYAFDTKOENLLKLAPDKFNDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLIA
 NFIELAPLNSNFIIQEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENIYKLIKID
 LLRLPSDIKHIFYGKX"
 6901..7941
 /gene="neuB1"
 6901..7941
 /gene="neuB1"
 /codon_start=1
 /transl_table=11
 /product="putative sialic acid synthase"
 /protein_id="AAK91726.1"
 /db_xref="GI:15193222"
 /translation="MKEIKIQNIISSEKAPLVVPEIGINHNSLELAKIMVDAAFSA
 GAKIKHQTHIVEDEMSKAARKVIPGNAKISYIEMOKCALDYKDELALKEYTEKGLI
 VYLSTPFRAGANRLEDWGVSAFKIGSGECNNYPLIKHIAAFKPMIVSTGMNSTESI
 KPIVKILLDNEIPIFLMHTTNLYPTPHNLVRLNMLLEKKFSCMGLSDHTONLAC
 LGAVVLGACVLERHFTDSMRHSGDIVCSMDTKALKKELIIQSEQAIIIRGNNEKKA
 KQEQVTIDFAFASVYSIKDKKGEVLSDMTWVKRPRGLGSGISAAEFENILKKALRDI
 ENDAQLSYEDFA"
 7938..9056
 /gene="neuC1"
 7938..9056
 /gene="neuC1"
 /codon_start=1
 /transl_table=11
 /product="putative N-acetylglucosamine-6-phosphate

2-epimerase"
/protein_id="AAK91727.1"
/db_xref="GI:15193223"
/translation="MKKILFITGSRADYSKIISLMTVRVNSSEFELYIYATGMHLSKN
FGYVKELYKNGFNKILVFINDKYIOTDKALATIDGFSRYANEIKLPDILVVGDR
EPLAAIVGALNNILVAHIEGEGISGTDIDSLRAHISLAHILHVLNDEFAKRLMG
EDEKSFITGSPDLELNDKISLSAKRYIDINENYALLMFPVTTETTSKNOAD
NLVKALIQSNKIVYIPNNDLGFELILOSIEBEFKNPRKLPFPSPRFEYFTLLKNA
DLFGNSCILKEALVLTAGILVGSQRNGNENTLKVNSDEILKRAINTIHKQ
DIFSACLEILDSKLFYEQLSQDFKLSQKVKDK"
9053..9718
/gene="neu1"
9053..9718
/gene="neu1"
/codon_start=1
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAK91728.1"
/db_xref="GI:15193224"
/translation="MSLAIPARGSGKIKNNLVLLNKKPLIYYTIKAALNAKSISK
VVSSDSDELNAVQSNQVILKRPISLAQDDPTSDKVLHALKFYKDYEDVVELOPT
SLRTWHINEAFNLYKSNANALISVSECNDKILKAFVNCNCGDLAGINCDEYFMP
RKLPTYMSNGAIYILKREFLNPFLOSQTKHFLMDESSLDIDCLDLAKVQEI
WRK"
9706..9900
/note="orf11; premature translational stop codon due to
variation in homopolymeric G-tract"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAK91729.1"
/db_xref="GI:15193225"

CDS

Query Match 2.8%; Score 82.4; DB 1; Length 11442;
Best Local Similarity 51.7%; Pred. No. 0.00037;
Matches 213; Conservative 0; Mismatches 196; Indels 3; Gaps 1;
QY 1296 AAGGTAACCTTACATCTATAGAAAGCTTTACCAATAGAGATTCACATATTCATAGAAT 1355
DB 11399 AATTTCTTAACCTTGATAGATTGTTTTATATTTTACAAAAATTTAGGAAAAATAT 11340
QY 1356 ACCTTTAGTTCTTATATATATCCCGCTTATAACTGTGCAAAATATATATCAAGATGTGT 1415
DB 11339 GCCACAACTTCTATCATATAATCCCGCTTTTAAATCTTGCATTATTTATCAAGGGTTT 11280
QY 1416 AGATAGTCTCTTAATCAAACTGTTCGATCTCGAGTTTGTATTTGTTACGATGGTTC 1475
DB 11279 GCAAGTTGTATAAATCAGACTTTAAAGATATTGAAATTTTGATTATCGATGATAAAG 11220
QY 1476 AACAGATAATACCTTAGAAGTGATCAATAGCTTTATGGTAATAATCCTAGGGTACGCAT 1535
DB 11219 TAAGATAATAGTTTAAATATGTTTGTAGATTGCTTAAAGATCAAGGATAAAAT 11160
QY 1536 CAT---GTCTAAACCAATCGCGGAATAGCCTCAGCATCAAAATCAGCCGTTCTTTTTCG 1592
DB 11159 CTTTCAAAATGAGAAAAATTTAGCAGCTTTTGCAAGTAGAAATTTGGGGGTTTGCATTC 11100
QY 1593 TAAGGTTATACATCGGCGAGTTAGATTTCAGATGATTATCTTGAGCTGACGCTGAGTTGA 1652
DB 11099 TAGTCTGTGATTTATATATGTTTATAGATAGTATGATTTTAAACACCTGATGCTTGTA 11040
QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTGTGCTTTTATACC 1704
DB 11039 AATAACATTTAAAGAGATGAAAAAGGTTTGTGTTGCTTTGTATGCC 10988

RESULT 12

AY044868/C

LOCUS

DEFINITION

AY044868 11455 bp DNA linear BCT 03-JAN-2002
Campylobacter jejuni heptosyltransferase I (waaC), putative
cgs; lipid A biosynthesis acyltransferase (htrB), putative
two-domain glycosyltransferase, putative glycosyltransferase,
putative beta-1,4-N-acetylgalactosaminyltransferase (cgtA),

putative beta-1,3-galactosyltransferase (cgtB),
alpha-2,3-alpha-2,8-sialyltransferase (cstII), putative sialic
acid synthase (neuB), putative N-acetylglucosamine-6-phosphate
(neuCI), CMP-Neu5Ac synthetase (neuAI), putative acetyltransferase,
and putative glycosyltransferase (waaV) genes, complete cds; and
heptosyltransferase II (waaF) gene, partial cds.
AY044868
VERSION 1 GI:15430476
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 11455
/organism="Campylobacter jejuni"
/strain="ATCC 43460"
/db_xref="ATCC:43460"
/db_xref="taxon:197"
/note="Penner-type strain O:41"
<1..348
/gene="waaC"
<1..348
/gene="waaC"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAK95995.1"
/db_xref="GI:15430477"
/translation="FAKEVNLSCIDETHIELAPKFSLEELMAETKMDLIIGNDSP
THLAPLNKASITIFGATPSYRNAFTHINKIIDAGKKIQNTKHKIDSKDFCITRIEE
DIFKLAKGLLNEK"
338..1225
/gene="htrB"
338..1225
/gene="htrB"
/codon_start=1
/transl_table=11
/product="lipid A biosynthesis acyltransferase"
/protein_id="AAK95996.1"
/db_xref="GI:15430478"
/translation="MKNSDRILSLYILKEFVTMPCEILHFLALIVARITFYLNNK
HRKLIIDNLQICFPYQTKERDKLSLIYENFAQFGIDCLQNTKREILNKVNFN
ENLLIDALQKRPITFTAHYGNWEILSLAYAAKYGAIISIVGKKLSIEMYLEISQR
TQFDIELDKKGIRQMSALKERALGILTDQCVENESVRLKFFNKENVYQMGASL
IAQSNALIIIPVAYKEDGKFCIEFFNAKDSQNASLELTLYQAQSCSEMICKRPWEY
FFHRRFASVDKEIYKGA"
1222..2769
/note="orf3"
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAK95997.1"
/db_xref="GI:15430479"
/translation="MNLKQISVIIIVNAQOTLLECLNSLKDFDEIILLNNESDNTL
KIANEFKNFANLHYYSATFGALKNLASYAKNDWILSIDADEVLENCIEKLK
LELQEDNIALSRKNLYKGWKACGMPDYLRFINKNFTRFNDSLVHESLVLPGNV
KKIHLKNGLRHYAFRDISHLINKMQYSSSLWAKQNIHKKSGVLKANLRAFPTFFRNY

Location/Qualifiers

```

LKNGLYCYGSGTTSVNCALGTEPKYMKLYELOKPKTKCTCALIITTYNOKERKLVLID
SVKNLTFPLNEVLADGSKEDTARLIEEVQKDPSPCSLKHHTWQDEGFRAAKSNKAI
RVKSEYIILIDGMILEKDFIRKHFEFARKKVILQSGRIILNKESEELIKNDNYSI
AFSKDFKYSFLAKIYIKISIDQKFTNKELIRGSKTCNNSFFKTFDKLUDGEN
ENFTGWRGDESEFVARELNFNGIPRRLKRAIAVHYHEENSCKMLESNHQIYLDITIK
NKKTSWR"
2773..3966
/note="orf4"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAK95998.1"
/db_xref="GI:15430480"
/translation="MKKIGVVPIYVNVKYLRECLDSVINOTYTNLEIMLVNDGSTDE
NSLIANKYITUKRITLFDKNGGLSARNVGIIEYSGKYKLLKNTQHKHENSILIEF
QLDGNPNYIKYKSOAFNNEKDLNTFTYPSIDYIIFLDSNHWELNCLIECVIRM
KNVGLVHDCTYEDNIKKNKHTRMIEIFNFKEECIITPREYANRALSVGSRDLSFG
WNGMIDNFLKQIKLFINIIFEDIHFIIFASANKIYVLSQKLYLCRLRANSISN
HDKITRANVSEYFKDIYETEGENAKAKNYLKAASRVITALKLIEFFKDKOSENALA
IKETFLPCYAKKALMIKKFKKDLNLKQLVLIRPFIQTKLPYDIWKFQWQIKKIFNN
KNIKN"
4015..5058
/gene="cgtA"
CDS
4015..5058
/gene="cgtA"
/codon_start=1
/transl_table=11
/product="putative
beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAK95999.1"
/db_xref="GI:15430481"
/translation="MLFOSYFVFIICLPFRKIRHKIKTFLKNIQORDKIDSYPK
KTLAQINKYNEDLKIINKALIGEHEKGYFNVDKSKDPKPLNPAWLRVKNEAITL
KASLESIIPATQGVIGVNDCTDGSSEILIECKQYPSFIPKYPYEIQIONPKSEEN
KLYSYNVASIPKDEWLKIDVDHIIYDAKLYKSFITPRKNKYDVVSYSRVYDIHFN
DNFFLCKDNNNGILDEPGDCLLNNYNLKWLEVLDIRNNNNKKATQKQSFSSNIHSE
QLKYKRRILFTELNNYHFFPLKHKRAQDIYKNNWISIEEFKKFYQLONINHKIEPSI
SKETLKFIPLF"
complement(5044..5949)
/gene="cgtB"
CDS
complement(5044..5949)
/gene="cgtB"
/codon_start=1
/transl_table=11
/product="putative beta-1,3-galactosyltransferase"
/protein_id="AAK96000.1"
/db_xref="GI:15430482"
/translation="MFKSIIPLTVNVEQYIARATESCINQTFKDIEIILVDDCGNDN
SINIAKESKDKRIKIIHNEKNLGLLRARYEGVKVANSPIYIMFLDIDDYLELNACEE
CIRILDEQDEVLDFVNAIVESNVISYKRFDFNSGFYKKEFVKKIIANKNLYTMWG
KLIRKLYLEAFASLRLEKDVKNMAEDVLLYPMLSQAQKIAYMNCNLYHYVPNNNS
ICNTNEVLVKNQIOELQVLVNLQRQYILNKYCSVLVYLKYLILYIQIYKIKRRTKLM
VILLAKINILTLKILFKYKKFLKQC"
6038..6913
/gene="cstII"
CDS
6038..6913
/gene="cstII"
/codon_start=1
/transl_table=11
/product="alpha-2,3-/alpha-2,8-sialyltransferase"
/protein_id="AAK96001.1"
/db_xref="GI:15430483"
/translation="MKKVIISGNPSLKEIDYSRLPNDFVRCNCFYEDKYKYLKKK
KRAVFNPSLFFEQYVTLKHLIQNEYELEIMCSNFNQAHLNFKVFTFYDFPDA
HLYGDFEKOLKEFVYKPHIEYFNORITSGVYMCVVAIALGYKEIYLSGIDFYQNGS
SYAPDFKQKNLLKLAPFNKDNDSHYIGHSKNTDIIKALEFLEKTYEIKLYCLCPNSLLA
NLELAPNLSNFIIEQKNNYTKDILIPSSEAGYKTKNINFKKIKENIYKYLKID
NLRPSDIDKHYFKG"
6914..7954
/gene="neuB1"
CDS
6914..7954
/gene="neuB1"
/codon_start=1

```

```

/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAK96002.1"
/db_xref="GI:15430484"
/translation="MKEIKIQNIILISEKAPLVVPEIGINHNHSGLELAKIMVDAAFSA
GAKLIKHOTHIVEDEMSAAKAKVIIPGNKAKISIEIYEMQKALNYKDELAKVEYTELGL
VYLSTPFSRAGANLEDMGVSAFKTSGECNNYPLIRKHAIAFAKFKKFWIVSTGNNSTESI
KPTVKILLDNEIPFVLMHTNLNLYTPHNLRLNAMLELKKFAKFSMWGLSDHTDNLAC
LGAVALGACVLERHFTDSMHRSGPDIVCSMDTQALKELIIQSEQMAIRGNNESSKAA
KQEOVTIDFAFASVVSIRKDKKEVLSMDNIWVRPGLGGISASEFENILGKRLARDI
ENDTQLSVEDET"
7951..9069
/gene="neuCl"
CDS
7951..9069
/gene="neuCl"
/codon_start=1
/transl_table=11
/product="putative N-acetylglucosamine-6-phosphate"
/protein_id="AAK96003.1"
/db_xref="GI:15430485"
/translation="MKKILFITGTRADYSKISLMYRVONSSEFELYTPATGMHLSKN
FGYTVKELYKNGFKNYIEFINYDKYQTDKALATTIDGFSRYVNEIKPDLIYVHGDRI
EPLAAIIVGALNNILVAHEGGEISGTIDDSLRLHAISKLAHLHLVNDPEAKRLMQLG
EDEKSIFIIGSPDLELNDNKISLSEAKKYDINYNVALLMHFVPTTEITSIKNOAD
NLVKALISNKNYIYIIPNNDLGFELLQSYEEELKNRPKLEPSLRFYFITLLKNA
DFIIGNSSCILKEALYLTAGILVSGRQNGRUGNENTLKVANSDEILKAKINTIHKKO
DLFSAKLEILNSSKLFFEYLLSGEFFKLNTQKVFKDIK"
9066..9731
/gene="neuA1"
CDS
9066..9731
/gene="neuA1"
/codon_start=1
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAK96004.1"
/db_xref="GI:15430486"
/translation="MSLAIIPARGSGSKIGKNLNLNKNPLIYTIKALNTKSIKSK
VYVSSDSILNVAKSONVDILKRPISLAQDDTSPKVLHLKFKYKDYKDVVFLQPT
SPRTSDNIHIDEAFNLTKSNANALISVSECDNKILKAFVNDYDGLAGICNDYEPFIP
ROKLPKTYMSNGAIIYLIKTEFLNPNPFLQNKTKHFLMDESSSLDIDCLDLKKVEQI
WRK"
9719..9910
/note="orf11; premature translation stop due to deletion
in homopolymeric G-tract"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAK96005.1"
/db_xref="GI:15430487"

```

```

Query Match      2.8%; Score 82.4; DB 1; Length 11455;
Best Local Similarity 51.7%; Pred. No. 0.00037;
Matches 213; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

QY 1296 AARGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTACATATTCATAGAAT 1355
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11409 AAATTTCTTAACTTTCATAGAATGTGTTTATTTATTTATTTACAAAATTTAGGAAAATAT 11350
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1356 ACCTTTAGTTTCTATTTATATCCCGCTTATAACTGTGCAAAATTTATTTCAAGATGTGT 1415
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11349 GCCACAACCTTCTATCATATATCCCGCTTTTAAATTTCTGTGATTTTATCTTAAAGAGCTTT 11290
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1416 AGATAGTGCCTTAACTAACTGTGTGCGATCTCGAGTTTGTATTTCCTAACGATGGTTC 1475
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11289 ACAAGTTGTATAATCAGACTTTAAAGATATTGAATTTTGATCATCATGATAAAG 11230
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1476 AACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGGTATAATCTCCTAGGGTACGAT 1535
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11229 TAAGATAATAATAGTTTAATAATGGTTTGTAGAAATTTGCTAAAAAAGATCCAAAGATAAAAAT 11170
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1536 CAT---GTCTAAACCAATGCGGAATAGCCTTCAGCATCAATGCAGCGGTTCCTTTTGC 1592
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11169 CTTTCAAAATGAAGAAAATTTAGGCACCTTTTGCAAGTAGAAAATTTGGGGGTTTTTCATTC 11110
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
/db_xref="GI:3724297"
/translation="MHKPTISSVIALTLTLLCGGSGNSGNTTPPVKYKNVNSFLDLDN
IAINTQNCQIFGYNEDOTKIVAVESAPENAVQIVVHDEHGNFVRYVNSIETNGLS
STFRQSDVPQNGYVSVYSEGRONFVSTYSKDFIPESFIYAKNKNNSNCLSGS
GNNSPQSDIYKSFHIPDENKLFPSLNFHKNLSPEPTNKLGRDVSVNSRRGRPLLL
TAYERTNQDYNLFSEFRNALSGQEIIDLIDQDRTPTWTVSDVEISSAQLF
VDGQFNAPYATLWMSLNTSSDKYSYSLNIDVNYLLKLGKEIKTSPSPYKFKQHV
SGTSSITSHLSAENITLDCFPNPEPTLGINDHNTQCIKITHQTKDNTILRLID
SELTVPSPGLPIFKOTIYSPYHNELPVMEKFDSPILDOELKEIKDFSVSLKIPSSSIT
ETFLYQDLYKRTIGSELIDFSIDNIPILKNISAOQDQDLLKROPYTWVLEE"
complement(10522..12714)
/gene="wbfb"
CDS
complement(10522..12714)
/gene="wbfb"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33587.1"
/db_xref="GI:3724298"
/translation="MKTGHRPLLNTSLSLGLVLISSALLTPROAFADFSYPLRHSSQ
TDFGCTGLMOPSAARMECELNLAITNDDYLHYAISLQLPWLTTVRYTQVHELL
YSDQASGDTKYTDKSIDAKLRLWQESYWLPELSGLRDIGCTGLFDEGLVATKQV
GPDFTLFGGWTGLGNRANLSDGKISSDCCGDSNNGGGMGLDGLRMTGCTLFGG
IEYQPLAPLRLQLEYNDRYDSDFVTRQCPDAMPVTPWNFGLTYALTDWADLRVS
YQGTNLTAGLQETGNLATLKPTWLDSPATYQPVPSKADLSDDEWQRLSQDVANLAG
YQVETLYQSDSVVYKGSQVYKRDQAEHRAATLLANTGINARYRIIETSDOQPLT
ETQLNSAEPKRIIDHDYPNARLSDAKQHTDPAPIHGDLKAARTERWKFGFSPVLOQSF
GAEDFYLYAIGNANASYKMGDHWLGSGLYGNLIDNDYDKNYTVPDGTNLKRVRT
LSRQTFEERLSHQLYTFDRWQHFTYQAGYLETMEFAGVGSSELYLPGLKNVAWG
IDGNYVQRPDSASGLYEHEHOYDVETGDNTRYQVGTTLTGHTATLYWQFQWLLDNT
LLKVSAGRYLTEDIGFTVDFSKQFAGVAGAFATKTDLSAEFGSGSTKGFYISIP
LDLMTVYSTERVNTSWLPLORDGGQMLSRQYGLYKMTDARSPWYERF"
complement(12711..13577)
/gene="wbfc"
CDS
complement(12711..13577)
/gene="wbfc"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33588.1"
/db_xref="GI:3724299"
/translation="MCQTDLPLRLSPFKICRSPKIKRAMSFITSRGWLTLAAAIISGH
SICVAAPLATSEVRYAHNNQAIHSTRLSFSEPMRLDSAVMQTLEQAGLNTQVWPS
AGLFDLSHAFLEKRDVLLKADQSSAPPTQOALWASLIAQLRQAEFAKRLFTISVDPD
WTRIAPOHNPRLNGSWLLTLNSKSTOVSVYGVNPGDVIWHNRLSAKDYAHAAGLID
EOISLTVIQPDGIAQKHAVYWNQDFNEVAPGAIYVVPPLKRAFFDFTVTDADLNQ
LVIELLRNREPL"
complement(13577..14266)
/gene="wbfd"
CDS
complement(13577..14266)
/gene="wbfd"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33589.1"
/db_xref="GI:3724300"
/translation="MHRLYLAFCOLFLIMGCSNSSHSETMQLAIFGTPAPVISPO
KVAGIPYSLYIQONDNPLALMWAEPTNRSRHRLAKWVSAGQEMLYTEGGRITKQ
NLNDGANLSLSDSDPLDGLHRASTPREWHFYLWQPGYHFRYLARHSQFYLIGEQ
QKOLPNGESRTLLRIDEQITIAALNQYHNOYWLPESTGAVTASEQOOLAPGTHRYALA
VAKSYLNQGGI"
14390..14638
/gene="orf6"
CDS
14390..14638
/gene="orf6"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33590.1"
/db_xref="GI:3724301"
/translation="MATTATAAAVTATAVVAVVPTAAVLAVEPALAAVSSVAVAKPL
TSSALSNTSAEFFIVISMKDIIYTDVHSPFRFRTCKR"
14935..15525
```

```
/gene="wbfe"
14935..15525
/gene="wbfe"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAA33591.1"
/db_xref="GI:3724302"
/translation="MQOSIHASLARNVFKTSPLRFLSTPLCOCVALMTSSLLTSSLLV
PSVAEPFACISLAYTNAEYQPNHAASTDGSPITAVALQGGTDFSDTLGLGEARTSTV
QRSSGVNIDNLSAAYKALNLPMSAQSLYGLAGYSYFSLDQKQGVSGVDESGFSGIGI
HYALDSRSAIAFDFFINHANGDEARLNTVNSQLKF"
15645..18380
/gene="wbfe"
15645..18380
/gene="wbff"
/note="probable transport of capsular precursors"
/codon_start=1
/transl_table=11
/protein_id="BAA33592.1"
/db_xref="GI:3724303"
/translation="MTQGYILRTLSSLVACLSPVAMAQTTPEQIQMFKNLPAAQQAQ
LASKYGFALPSATNSASQYDNPEVMLPQNLNENITAOPLNGODALERPLGLSLFSG
SPSTFAPGVDPVPSDYTVGAGDEIVQLFKQKQTHRLRYNRAGVINFTSLGVQVA
GLKFSEVRQSLGQVRVEQMGIRSDISIDELRTMQIFVMGDAYKPGATVTSALTLSQ
AIYISGFSQSGALNLTQLKREGVLQDLIDLKLDGNNRNDVRLAMPDGVVFTGAVD
NTEIIRGEVNRPAVYEVKSTETYQDIIKMGAGFTANAYROKTEVKRYASDGIRKALT
NLNEGLOTLKNGDAINVLKKTQALAHYVEIVGDVHOHGFIEWKFGLRVADVQSV
DSAFNATDNYVALVVRQKNLQRDIEVQOINLANALEPGSRDNLTLQSRDRLLIENR
FNFDYDNTDLSDKETENTAKTLEQAQOATQAKLEQKQOQINNTVRYNASALNPS
LPNNLSQLPSGQSLEAHPKAIQFAGQEMTLEAEFKLQNTNRLLAPILQLLYQ
QAQVGSQPIAEVYGVKYPYPIAAMPVSALLEAAGGLTFNAYTLRABLARREID
PNHRYVTQISSLDLROMLSEPSANRWTIOGMKLNVLKPAKIQOTVVLQGVLFPP
CTYVRQGETLAELLRAGLTPEFANPRGAIFREALRLQEQKLLNOYAADRAETAK
KTPRADNLSSVSDPAKTLAFVEASRALGRYVQLSRIMRDEPADFMLEDGDF
LYPTFRNTISINGEVQVPTYLLDSGLNVDYDLYNKAGGVKKQADARIFVVRADGSV
18485..19492
/gene="wzz"
18485..19492
/gene="wzz"
/note="probable regulation of O-antigen length"
/codon_start=1
/transl_table=11
/protein_id="BAA33593.1"
/db_xref="GI:3724304"
/translation="MQEQKUTPPNYLTPYPAQPTAIGDDEIDLRELFRALWKGKWTIIA
ITLVFAIGSVIFAIMQPNIIYKAEALIAPESEQGGLSALASQFGGLASLAVNLGGK
GYVDKTMQAEVLKSRQTFEFTQOHHILADIMAAKKWDSADQIYDPALYNAQTHT
WREAKAPFKPEPSMOCAYKVESKLHVNKAKDSGMVTIAVHLSPTVAQOMQWTWIE
DINKVMKRDVAEHRSTAELEQIALTNVADIFTVLYKLEEQAKTIFAEVRDEYV
FKTIDPALAPEAKPRALICVIGTILGMLGVAILVLRPAFRKAGKNHEHADVRET
A"
19650..20777
/gene="IS1358"
19650..20777
/gene="IS1358"
/note="probable insertion sequence like element"
/codon_start=1
/transl_table=11
/protein_id="BAA33594.1"
/db_xref="GI:3724305"
/translation="MSELINPEMFHFIKDYROFSKVEHKLSDIILLTCGLVLSCHDG
WDGIIDFGNARLDLFKRYGHFEAGIPSDATLSRVGMINPVALORSFIANKNCHTILT
DGEVIAIDGKTLRSGYDRSGKGTIHHVNAFATANGMSIQLKYSNEITAIPKLL
DLDDVKGLTIDAMGCKKIAQIRKADYLLAVKNGMQLQAEQAFDDYFRMDMLQD
FDGSSYSTQKSHGRIETRYALVRNLDLSVLGDIIEHPWGLKSMGIVASIROESAVATE
QDVSIRYICSKLEAQTLLAEATRSHMGVEVHWSLDTAFCDNSRTRADDAEAFAR
IROICLNLLKSETTFKGIKRRKNCAMDEKYLKSVLESLT"
20961..22079
/gene="gmd"
20961..22079
/gene="gmd"
```



```
LLSLAKTIPTIGYAIWGTIGAGVSVTLGMIVFKERKSVGKLLFTIMTIAGVVGKLK
TSGV"
join(2650. .3774,3788. .3822)
/ gene="ddla"
RBS
2650. .2655
/ gene="ddla"
CDS
2662. .3774
/ gene="ddla"
/codon_start=1
/transl_table=11
/product="D-alanine--D-alanine ligase"
/protein_id="CAC98933.1"
/db_xref="GI:16410258"
/db_xref="SPTREMBL:Q8Y8P1"
/translation="MKTKLILLYGKSAEHSVLSQAFSVINALDLKLEKFAAIIYITN
EGEITQGLSLKDFVEOLREFSATDTIKLATSESEKSEGEALSPAVLEADGOETVE
RLIDPNGEDSGTVOGLFEVLNIPDPYGVNGVNLASSAAMDKIVMKKIFADAGIQPQAVAV
RLIDMKNIQEEWVAEMEEVLTYPVFVKPNALGSSVIGISKATNKKKELVDAWTEAFLYDR
RVVVGQVVAIEHTEHGVLDNTPVCSVGEIPEGAVALFYDYKAKYQDNNTALIIPT
EYDPELEQMKBYAQAFGLDASGLVRADFLEDNQLFLNEVNTMPGFTPYSMYPL
LWQETGLPYGALIERLVDLAKERHAAKNALKYKLED"
3788. .3822
terminator
/ gene="ddla"
join(3822. .3828,3838. .5211)
/ gene="murF"
RBS
3822. .3828
/ gene="murF"
CDS
3838. .5211
/ gene="murF"
/codon_start=1
/transl_table=11
/product="UDP-N-acetyl-muramoylalanyl-D-glutamyl-2,
6-diamino p-melate-D-alanyl-D-alanyl ligase"
/protein_id="CAC98934.1"
/db_xref="GI:16410259"
/db_xref="SPTREMBL:Q8Y8P0"
/translation="MKKTGVEVVAMLDSSIHVDAFRDVVITGVCFDTRIKSGDLFVP
FYGNVRDGHFVGQAREMGAVATFWQKNVPPNPTDFPVLIVEDTLALQELAARYIQQ
VFKPVIAITGSGKTTTKDIMAIAIVETTVKHYVTGNGFNHHLGLPTVILTMPEDTEVA
VLEMGHNRHEITVLSKIAKPDIAITNTGEAHEIVLGSREEIAKAKLETACGLNPSG
ILLYHEETVLLGINDGFRQLTFCKSEAAEYPLLEIRAEAGTSFTWWEPELEIFV
PIIHEHNTHMAALAREIGIEGEKIQSALKNMERSKSRLEWITKSGARILNDAY
NYSPTALKTVLFTFMHDSGPKPLVLDLMLDELGLSLKHLGSAEVLENDTIQKVF
LXGDMAKAFGEVAESKIGQGKVHHFDTKESLETALLSEIKGNNEWILVKGSGYGMGLKDV
VENLIK"
join(5263. .5268,5278. .5991)
/ gene="lmo0857"
RBS
5263. .5268
/ gene="lmo0857"
CDS
5278. .5991
/ gene="lmo0857"
/ note="similar to carboxylesterase"
/codon_start=1
/transl_table=11
/protein_id="CAC98935.1"
/db_xref="GI:16410260"
/db_xref="SPTREMBL:Q8Y8N9"
/translation="WVACLCIHGFTGSPSEVKPLADYLRHETDWDVLAFLPGHDHLR
HLKNVYTKDWIVFVDSILSOMLKEDDEVYIIGFSGMGLLAGWLARHYPEVKKLVLLST
AVNMEHPTLVNSKQVLTEAKEVTLKNSPMFKRYQKKVETPTWTLQPKRKNVALAK
PVFEHIEIPTFAQGSADQVVPAAKSVNFMESIPGPKELFIEGSKRHVICQDEQADK
LFAAVLTLQKDVAVLNTK"
complement(join(5990. .7028,7035. .7040))
/ gene="lmo0858"
terminator
complement(5990. .6032)
/ gene="lmo0858"
RBS
5998. .6036
/ gene="lmo0857"
terminator
5998. .6036
/ gene="lmo0857"
complement(6003. .7028)
CDS
/ gene="lmo0858"
/ note="similar to transcription regulator lacI family"
```

```
/codon_start=1
/transl_table=11
/protein_id="CAC98936.1"
/db_xref="GI:16410261"
/db_xref="SPTREMBL:Q8Y8N8"
/translation="MKKTMODIADRLNITKNSVSQALGNKNKGVSEOTKFLAVFKMADE
LGYYKREIAREVVKERFALLATSFALSQSFSGEIIDNMKQSIHAHQAEIIFPVTD
EEAATNOLPQLTSENWTFGLSHINTDYSKVIDLGMVPLVLDHDPHLKADAVIS
ONKDGAPMAVEFLIONKHOKIGFLGDVTFSPSYERLEGYKKALQYHHPNKKYAIT
RIKEBOTLYRTLDELDELPTAFWFCVNSGLGLFINTVYLOSQGYNIPKDLSLIICPDNTE
FTVLSNPQLITMCTNLSFMGEKAVELMYNRIRKPDDEGFVHLALATNLILARDSVGENKK
Query Match 2.8% Score 82.4; DB 1; Length 270050;
Best Local Similarity 55.0%; Pred. No. 0.00017;
Matches 186; Conservative 0; Mismatches 146; Indels 6; Gaps 1;
Qy 1345 ATTCATAGAATACCTTTAGTTTCTTATTTATATCCCGCTTATAACTGTGCAAAATATATT 1404
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234478 ATTCTTTTGGAACTTTAGTAGTGTAAATTTATCCAGTTTATAATGTAGAAAATATGTT 234537
Qy 1405 CAAAGATGTGTAGTAGTGCTCTTAATCAAACTGTTGTCGAGTCTCGAGTTTGTATTCT 1464
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234538 AACGATGCTTGGATTCAGTTCTTTGAGCAACATATACCTCGAAGTTATCGTAGTT 234597
Qy 1465 AACGATGTTCAACAGATAATACCTTTAGAAGTGAATCAAGCTTTATGGTAATAATCCT 1524
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234598 AATGATGAGCGACTGATAATTACGTAAGTCATAAAAAGTATCTCAGATAA-----T 234651
Qy 1525 AGGTACGCATCATGCTCTAAACCAANTGGCGAATAGCCTCAGCATCAANTGCAGCCGTT 1584
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234652 CGAATTCCTTTTGTAAAAAGAAAATGCGGACAAGTACTGCCCGAAATTTTGGATT 234711
Qy 1585 TCCTTTGCTAAAGGTTATTACATTTGGGCAGTTTAGATTTCAGATGATTATCTTGAGCCTGAT 1644
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234712 GATGTGGCTACGGGAGATTACATCGTGATGTTAGTAGATTAGTAGATTACATTAGCAAAAAC 234771
Qy 1645 GCAGTTGAACGTGTGTTTAAAAAGAAATTTTAAAAAGATAA 1682
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234772 CTAGTAGAAACTTGTGTTGATACTCTACAAAAAACGAA 234809
```

Search completed: January 3, 2003, 23:56:07
Job time : 8776 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 20:10:07 ; Search time 574 Seconds
(without alignments)

11687.636 Million cell updates/sec

Title: US-09-842-484A-1

Perfect score: 2979

Sequence: 1 ttataaacgattataagaag.....acattgcattttattataaaa 2979

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2979	100.0	2979	21 AAA27449	P. multocida chond
2	2979	100.0	2979	24 ABA05097	Pasteurella multoc
3	2963	99.5	2979	24 ABA05098	Pasteurella multoc
4	2283.4	76.6	2937	20 AAX58857	Pasteurella multoc
5	2283.4	76.6	2937	21 AA235589	P. multocida hyalu
6	1503.2	50.5	2112	21 AAA27448	P. multocida hyalur
7	86	2.9	5550	20 AAX13074	Enterococcus faeca
8	81	2.7	984	24 ABQ68041	Listeria monocytog
9	81	2.7	984	24 ABQ69955	Listeria monocytog

10	80	2.7	19966	21	AAZ45258	DNA encoding enzym
11	80	2.7	19966	21	AAZ30355	Nucleotide sequenc
12	79.2	2.7	11474	21	AAAS3720	LPS core biosynthe
13	77.2	2.6	12237	24	ABL34358	Human immune syste
14	77.2	2.6	17527	24	ABL33432	Human immune syste
15	77.2	2.6	17527	24	AAAS6332	Chemically pretrea
16	76	2.6	7458	21	AAA70106	Plasmodium falcipa
17	75.6	2.5	963	24	ABN67436	Streptococcus poly
18	74.8	2.5	15548	24	ABL34155	Human immune syste
19	74.6	2.5	13807	18	AAAT62328	Exopolysaccharide
20	74.6	2.5	14602	18	AAAT73236	Streptococcus ther
21	74	2.5	40862	24	ABL34072	Human immune syste
22	73.2	2.5	6992	21	AAZ60930	Nucleotide sequenc
23	73	2.5	1082	24	ABQ69025	Listeria monocytog
24	73	2.5	1162	24	ABQ70597	Listeria monocytog
25	71.8	2.4	9106	24	ABL54369	Chemically treated
26	71.8	2.4	17280	22	AAAS46772	Tumour suppressor
27	70.8	2.4	981	24	ABN67435	Streptococcus poly
28	70.6	2.4	5748	24	ABL33142	Human immune syste
29	70	2.3	14924	24	ABL54321	Chemically treated
30	70	2.3	14924	24	ABL32224	Human immune syste
31	69.8	2.3	6048	24	ABQ67002	Human anglogenesis
32	69.8	2.3	16258	24	ABL70376	Chemically treated
33	69.8	2.3	16258	24	ABK40038	Human chemically p
34	69.4	2.3	7442	22	AAAS46686	Tumour suppressor
35	68.8	2.3	6109	24	ABL32326	Human immune syste
36	68.8	2.3	6109	24	AAAS61077	Human gene regulat
37	68.8	2.3	9095	24	ABQ67061	Human anglogenesis
38	68.8	2.3	12507	24	ABL32299	Human immune syste
39	68	2.3	5181	24	ABL70443	Chemically treated
40	68	2.3	5493	24	AAAS1386	Human gene regulat
41	68	2.3	6244	24	ABL32484	Human immune syste
42	68	2.3	15950	21	AAZ45260	DNA encoding enzym
43	68	2.3	15950	21	AAZ30357	Nucleotide sequenc
44	67.8	2.3	2325	24	ABQ67858	Listeria innocua D
45	67.8	2.3	2325	24	ABQ69470	Listeria innocua D

ALIGNMENTS

RESULT 1
AAA27449
ID AAA27449 standard; cDNA; 2979 BP.
XX
AC AAA27449;
XX
XX
DT 17-AUG-2000 (first entry)
XX
DE P. multocida chondroitin synthase coding sequence.
XX
KW Chondroitin synthase; CS; enzyme; hyaluronic acid; ulcer;
KW tissue abrasion; viscoelastic replacement; bioadhesive;
KW ss.
XX
OS Pasteurella multocida.
XX
FH Key Location/Qualifiers
FT CDS 61..2958
FT FT /*tag= a
FT FT /product= "PmCS"
XX
PN WO200027437-A2.
XX
PD 18-MAY-2000.
XX
PF 10-NOV-1999; 99WO-US26501.
XX
PR 11-NOV-1998; 98US-0107929.
PR 01-APR-1999; 99US-0283402.
XX
PA (OKLA) UNIV OKLAHOMA STATE.
XX

1 TTATAAAGGTAACGATTCAAGCAAGGTTATTTTAAAGGAAGAAA 60

|||||
Db 1 TTAATAACTGATTAAGAGGTAAACGATTCAAGCAAGTTAAATTTTAAAGGAAGAA 60
QY 61 ATGAATACATATTACAAAGCAATAAAAGCATATAACAGCAATAGCATATGAATAGCACTC 120
Db 61 ATGAATACATATTACAAAGCAATAAAAGCATATAACAGCAATAGCATATGAATAGCACTC 120
QY 121 AAATTTATTGAGAAGTCTGCTGAACCTTACGGCGCAAAATCGTTGAATTCCAAATATTC 180
Db 121 AAATTTATTGAGAAGTCTGCTGAACCTTACGGCGCAAAATCGTTGAATTCCAAATATTC 180
QY 181 AAATGTAAGAAAAAATCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGTT 240
Db 181 AAATGTAAGAAAAAATCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAACAGTGTT 240
QY 241 TCGGATAGCTCATATAGATATCCACACAGCTCTTACTTCCAAAGTAAAAAAATTAAT 300
Db 241 TCGGATAGCTCATATAGATATCCACACAGCTCTTACTTCCAAAGTAAAAAAATTAAT 300
QY 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTATCACTGGGAAAAAA 360
Db 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGGAATCTATCACTGGGAAAAAA 360
QY 361 TCGGAGAACCGCAAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTTAAAGATCTT 420
Db 361 TCGGAGAACCGCAAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTTAAAGATCTT 420
QY 421 GTTCTTGCTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA 480
Db 421 GTTCTTGCTCCATTGCCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA 480
QY 481 AGCTTAGGTATAAGGCTGTAATAGAAATATCGGTCTTTCTATTATTATTCCTACATTT 540
Db 481 AGCTTAGGTATAAGGCTGTAATAGAAATATCGGTCTTTCTATTATTATTCCTACATTT 540
QY 541 AATCGTAGCCGATTTTAGATATAACGTTAGCGTGTGGTCAATCAGAAAAACAACCTAC 600
Db 541 AATCGTAGCCGATTTTAGATATAACGTTAGCGTGTGGTCAATCAGAAAAACAACCTAC 600
QY 601 CCATTTGAAGTCTGCTGCAGATGATGTTAGTAAAGAAAAAATCTTACCTAGTTGCGAA 660
Db 601 CCATTTGAAGTCTGCTGCAGATGATGTTAGTAAAGAAAAAATCTTACCTAGTTGCGAA 660
QY 661 AAATAGCAACAAAACTTGACATAAAGTATGTAAGCAAAAAAGATTATGGATATCAATTG 720
Db 661 AAATAGCAACAAAACTTGACATAAAGTATGTAAGCAAAAAAGATTATGGATATCAATTG 720
QY 721 TGTGCAGTCAGAACTTAGGTTTACGTACAGCAAGTATGATTTTGTCTCGATTCTAGAC 780
Db 721 TGTGCAGTCAGAACTTAGGTTTACGTACAGCAAGTATGATTTTGTCTCGATTCTAGAC 780
QY 781 TCGGATATGSCACCAACAATATATGGTTCTATTCTTACAGAACTATTAGAAAGAC 840
Db 781 TCGGATATGSCACCAACAATATATGGTTCTATTCTTACAGAACTATTAGAAAGAC 840
QY 841 AATGATATTGTTTAAATGGACCTAGAAAAATATGTGGATACTCATATATATACCGCAGAA 900
Db 841 AATGATATTGTTTAAATGGACCTAGAAAAATATGTGGATACTCATATATATACCGCAGAA 900
QY 901 CAATTCCTTAACGATCCATATTTAATAGAACTACTACCTGAAACCGCTACAAATAACAA 960
Db 901 CAATTCCTTAACGATCCATATTTAATAGAACTACTACCTGAAACCGCTACAAATAACAA 960
QY 961 CCTTCGATTACATCAAAAGGAATATATCGTTGGATTGGAGATTAGAACTATTCAAAAAA 1020
Db 961 CCTTCGATTACATCAAAAGGAATATATCGTTGGATTGGAGATTAGAACTATTCAAAAAA 1020
QY 1021 ACCGATAATCTACGCTATGATCTCCGTTTCGTTTATTTGTTGCGGGAATGTTGCA 1080
Db 1021 ACCGATAATCTACGCTATGATCTCCGTTTCGTTTATTTGTTGCGGGAATGTTGCA 1080
QY 1081 TTTTCTAAAGAAAGGCTTAATAAGTAGTGTGTTTCCATGAGAAATTTAATCATTTGGGG 1140
|||||

Db 1081 TTTTCTAAAGAAAGGCTTAATAAGTAGTGTGTTCCGATGAAGAAATTTAATCATTTGGGG 1140
QY 1141 GCGGAAGATGTAGAATTTGGTTACAGATTATTTGCCAAAGCGCTGTTTTTCAGAGTAAT 1200
Db 1141 GCGGAAGATGTAGAATTTGGTTACAGATTATTTGCCAAAGCGCTGTTTTTCAGAGTAAT 1200
QY 1201 GACGCGGAATGGCCATCCATCAAGAACCCCTGTTAAAGAAAAATGAAACAGAACCGAA 1260
Db 1201 GACGCGGAATGGCCATCCATCAAGAACCCCTGTTAAAGAAAAATGAAACAGAACCGAA 1260
QY 1261 GCTGTTAAAGTATTACGCTTAAATTTGTAAGAAAAAGTACCTTACATCTATAGAAAG 1320
Db 1261 GCTGTTAAAGTATTACGCTTAAATTTGTAAGAAAAAGTACCTTACATCTATAGAAAG 1320
QY 1321 CTTTACCATAAGAGATTACATATTTACATAATACCTTTAGTTCTTATTTATATATCC 1380
Db 1321 CTTTACCATAAGAGATTACATATTTACATAATACCTTTAGTTCTTATTTATATATCC 1380
QY 1381 GCTTATAACTGTGCAAAATTTATTTCAAAGATGTGTAGATAGTCTTAAATCAAACTGTT 1440
Db 1381 GCTTATAACTGTGCAAAATTTATTTCAAAGATGTGTAGATAGTCTTAAATCAAACTGTT 1440
QY 1441 GTGATCTCGAGTTTGTATTTGTAACGATGTTTCAACAGATAATACCTTTAGAAAGTATC 1500
Db 1441 GTGATCTCGAGTTTGTATTTGTAACGATGTTTCAACAGATAATACCTTTAGAAAGTATC 1500
QY 1501 AATAAGCTTTATGGTAAATATCTTAGGGTACGATCATGTCTAAACCAATGCGGAAAT 1560
Db 1501 AATAAGCTTTATGGTAAATATCTTAGGGTACGATCATGTCTAAACCAATGCGGAAAT 1560
QY 1561 GCTCAGCATCAATCGACCGCTTTCTTTGCTTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
Db 1561 GCTCAGCATCAATCGACCGCTTTCTTTGCTTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
QY 1621 TCAGATGATTATCTTGAGCTGATGCACTGTTGAACCTGTGTTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATTATCTTGAGCTGATGCACTGTTGAACCTGTGTTTAAAGAAATTTTAAAGAT 1680
QY 1681 AAAACGCTAGCTTGTGTTTATACACTAATAGAAACGTCATCCGGATGCTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTTGTGTTTATACACTAATAGAAACGTCATCCGGATGCTAGCTTAATC 1740
QY 1741 GCTTAATGGTTTACAAATGGCCAGAAATTTTACAGAGAAAACTCACAACGGCTATGATTGCT 1800
Db 1741 GCTTAATGGTTTACAAATGGCCAGAAATTTTACAGAGAAAACTCACAACGGCTATGATTGCT 1800
QY 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTAACGGATGATTAAAGCAAAAT 1860
Db 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTAACGGATGATTAAAGCAAAAT 1860
QY 1861 ATTGAAACCGCGTGGAATTTATGACATGTTCCCTTAAACTCAGTGAAGTTGGAATTTTAA 1920
Db 1861 ATTGAAACCGCGTGGAATTTATGACATGTTCCCTTAAACTCAGTGAAGTTGGAATTTTAA 1920
QY 1921 CATCTTAATAAATCTGCTATAACCGGTTTACATGGTGATAACACATCCATTTAAAGAA 1980
Db 1921 CATCTTAATAAATCTGCTATAACCGGTTTACATGGTGATAACACATCCATTTAAAGAA 1980
QY 1981 CTGCGCATTCAAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGC 2040
Db 1981 CTGCGCATTCAAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGC 2040
QY 2041 ATCAATTTATTAATTTATGACAAATTTTGAATGATTTAGTGAAGTAGAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATTTATGACAAATTTTGAATGATTTAGTGAAGTAGAAGTATATCTTC 2100
QY 2101 AATAAAACCGCTGAAATCAAGAAAGAAATGATATGTTTAAAGAGTCTTAAACTCATTC 2160
Db 2101 AATAAAACCGCTGAAATCAAGAAAGAAATGATATGTTTAAAGAGTCTTAAACTCATTC 2160
QY 2161 AATAAAGATGCAAAATCGCAGTCAAGTATTTTCTATCCCAATACATTTAAACGGCTTAGT 2220
Db 2161 AATAAAGATGCAAAATCGCAGTCAAGTATTTTCTATCCCAATACATTTAAACGGCTTAGT 2220

```
Qy 2221 AAAAACTAAACAAATATTATTGAATATAATAAATAATATATTCGTTATTATTCTACATGTT 2280
Dy 2221 AAAAACTAAACAAATATTATTGAATATAATAAATAATATATTCGTTATTATTCTACATGTT 2280
Qy 2281 GATAAGAAATCATCTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAGCAC 2340
Dy 2281 GATAAGAAATCATCTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAGCAC 2340
Qy 2341 CAAGTGAATATTTACTAAATAAATAGACATCTCATATTACAGGAGTAATAGACTAATAAAA 2400
Dy 2341 CAAGTGAATATTTACTAAATAAATAGACATCTCATATTACAGGAGTAATAGACTAATAAAA 2400
Qy 2401 ACTGAGGCACATTTAAAGTAATTAATAAATAAGCTAGTTAAATCTAAATTTGTGAATAC 2460
Dy 2401 ACTGAGGCACATTTAAAGTAATTAATAAATAAGCTAGTTAAATCTAAATTTGTGAATAC 2460
Qy 2461 ATCATTTTGTATATCATGACAGCCCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Dy 2461 ATCATTTTGTATATCATGACAGCCCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Qy 2521 AAATATGATGTCGCATCAATTTCTCAGCATTAACACATGATTGTCGAGAAAAATCAAT 2580
Dy 2521 AAATATGATGTCGCATCAATTTCTCAGCATTAACACATGATTGTCGAGAAAAATCAAT 2580
Qy 2581 GCGCATCCACCATTTAAAAAGCTGATTAACACCTATTTTAATGACAATGACTTAAGAAGT 2640
Dy 2581 GCGCATCCACCATTTAAAAAGCTGATTAACACCTATTTTAATGACAATGACTTAAGAAGT 2640
Qy 2641 ATGAATGTGAAGGGGCATCAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Dy 2641 ATGAATGTGAAGGGGCATCAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Qy 2701 CTGAGGATTAATAAGAAGTATCATACATCCTGCCAATCAATGATAGTGTGCCAAGATAT 2760
Dy 2701 CTGAGGATTAATAAGAAGTATCATACATCCTGCCAATCAATGATAGTGTGCCAAGATAT 2760
Qy 2761 AACACTGAGGATATTGCTTCCAAATTTGCACTTTTAACTTTAGAAAAAGAAACCGGCAT 2820
Dy 2761 AACACTGAGGATATTGCTTCCAAATTTGCACTTTTAACTTTAGAAAAAGAAACCGGCAT 2820
Qy 2821 GTATTTAATAAACATCGACCCCTGACTTATATGCTTGGGAACGAAAAATTAACAATGGACA 2880
Dy 2821 GTATTTAATAAACATCGACCCCTGACTTATATGCTTGGGAACGAAAAATTAACAATGGACA 2880
Qy 2881 AATGAACAAATTCAAAGTGCRAAAAAAGCGGAAAAATATCCCGTTTAAACAAGTTCAATTAT 2940
Dy 2881 AATGAACAAATTCAAAGTGCRAAAAAAGCGGAAAAATATCCCGTTTAAACAAGTTCAATTAT 2940
Qy 2941 AATAGTATAAGCTATAAAACATTTTGCATTTTATTAAAAA 2979
Dy 2941 AATAGTATAAGCTATAAAACATTTTGCATTTTATTAAAAA 2979
```

RESULT 3

ABA05098

ID ABA05098 standard; DNA; 2979 Bp.

XX AC ABA05098;

XX DT 22-FEB-2002 (first entry)

XX DE Pasteurella multocida chondroitin synthase gene #2.

XX KW Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;

XX KW eye application; joint application; moisturiser; drug delivery;

XX KW wound dressing; biocompatible film; ds.

XX OS Pasteurella multocida.

XX PH key

XX CDS 61..2958

XX FT /*tag= a

/product= "chondroitin synthase"

WO200180810-A2.

01-NOV-2001:

25-APR-2001; 2001WO-US13395.

25-APR-2000; 2000US-199538P.

(DANG/) DE ANGELIS P L.

De Angelis PL;

WPI; 2002-049237/06.

P-PSDB; AAM47336.

New chondroitin synthase gene obtained from Pasteurella multocida, useful as hyaluronan polysaccharide substitute in medical or cosmetic applications, e.g. for eye or joint applications, for moisturizer or wound dressings

Claim 4; Page 120-121; 125pp; English.

The present invention relates to the coding sequence of the Pasteurella multocida chondroitin synthase. A chondroitin polysaccharide may be used as a hyaluronan polysaccharide substitute in medical or cosmetic applications, for example in eye or joint applications, for moisturiser or wound dressings. The enzyme may be used in covalently coupling specific drugs, proteins or toxins to the structurally modified chondroitin for general or targeted drug delivery or radiological procedures, covalently cross linking the hyaluronic acid itself or to other supports to achieve a gel or other three dimensional biomaterial with stronger physical properties, and covalently linking hyaluronic acid to a surface to create a biocompatible film or monolayer. The present sequence is one version of the coding sequence of the invention.

SQ Sequence 2979 Bp; 1130 A; 466 C; 495 G; 888 T; 0 other;

Query Match

Best Local Similarity 99.5%; Score 2963; DB 24; Length 2979;

Matches 2969; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TTATAAATCGATTAAAGAAGGTAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA 60

Dy 1 TTATAAATCGATTAAAGAAGGTAACGATTCAAGCAAGGTTAATTTTAAAGGAAGAAA 60

Qy 61 ATGAATACATTTATCACAAGCAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120

Dy 61 ATGAATACATTTATCACAAGCAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120

Qy 121 AAATTTATTTGAGAAAGTCTGCTGAAACCTAGCGGCAAAATCGTTGAATTTCCAAATATTC 180

Dy 121 AAATTTATTTGAGAAAGTCTGCTGAAACCTAGCGGCAAAATCGTTGAATTTCCAAATATTC 180

Qy 181 AAATGTAAGAAAAAATCTCGACCAATTTCTATGTAAGTGAAGTAAAAAATATCTACTGTT 240

Dy 181 AAATGTAAGAAAAAATCTCGACCAATTTCTATGTAAGTGAAGTAAAAAATATCTACTGTT 240

Qy 241 TCGATAGTCTATTAGATATCGCAACACAGCTCTTACTTTCCAAACGTAATAAATAATTAAC 300

Dy 241 TCGATAGTCTATTAGATATCGCAACACAGCTCTTACTTTCCAAACGTAATAAATAATTAAC 300

Qy 301 CTATCCGAATCAGAAAAAATAGTTTAAAAAATAAATGGAATCTATCTACTGGAAGAAA 360

Dy 301 CTATCCGAATCAGAAAAAATAGTTTAAAAAATAAATGGAATCTATCTACTGGAAGAAA 360

Qy 361 TCGGAGAACGCAAGAAATCAGAAAGTGGAACTAGTACCCCAAGATTTTCTTAAGATCTT 420

Dy 361 TCGGAGAACGCAAGAAATCAGAAAGTGGAACTAGTACCCCAAGATTTTCTTAAGATCTT 420

Qy 421 GTTCTTGTCTCCATGTCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAA 480

Dy 421 GTTCTTGTCTCCATGTCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAA 480

Db 421 GTTCTTGCTCCATTGCCAGATCATGTTAATGAATTTTACATGGTACAAAAATCGAAAAA 480
Qy 481 AGCTTAGGTATTAAGCCTGTAAATAGAAATATCGGTCCTTCTATTAATTTTCTACATTT 540
Db 481 AGCTTAGGTATTAAGCCTGTAAATAGAAATATCGGTCCTTCTATTAATTTTCTACATTT 540
Qy 541 AATCGTAGCGGTATTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACACTAC 600
Db 541 AATCGTAGCGGTATTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACACTAC 600
Qy 601 CCATTTGAAGTCGTTTGCAGATGATGGTAGGTAAGAAAAAATCTTACCATTGTGCAA 660
Db 601 CCATTTGAAGTCGTTTGCAGATGATGGTAGGTAAGAAAAAATCTTACCATTGTGCAA 660
Qy 661 AATAGCAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG 720
Db 661 AATAGCAACAAAACTTGACATAAAGTATGTAAGACAAAAAGATTATGGATATCAATTG 720
Qy 721 TGTGCAGTCAGAACTTAGGTTTACGTACAGCAAAAGTATGATTTGTCTCGATTCTAGAC 780
Db 721 TGTGCAGTCAGAACTTAGGTTTACGTACAGCAAAAGTATGATTTGTCTCGATTCTAGAC 780
Qy 781 TGCAGATATGSCACCAACAATTTATGGTTTCATCTTATCTTACAGAACTATTAGAAAC 840
Db 781 TGCAGATATGSCACCAACAATTTATGGTTTCATCTTATCTTACAGAACTATTAGAAAC 840
Qy 841 AATGATATTTGTTTAAATGACCTAGAAAATATGGATATCTCATATAATTACCGCAGAA 900
Db 841 AATGATATTTGTTTAAATGACCTAGAAAATATGGATATCTCATATAATTACCGCAGAA 900
Qy 901 CAATTCCTTAACGATCCATATTTAATAGAACTCACTACTGTAACCGCTACAAAATAAACAAT 960
Db 901 CAATTCCTTAACGATCCATATTTAATAGAACTCACTACTGTAACCGCTACAAAATAAACAAT 960
Qy 961 CCTTCGATACATCAAAAGGAATATATCGTTGGATTGGAGATTAGAACTTTCAAAAA 1020
Db 961 CCTTCGATACATCAAAAGGAATATATCGTTGGATTGGAGATTAGAACTTTCAAAAA 1020
Qy 1021 ACCGATATCTACGTCATGATGATTCCTCGTTTCTGTTTATTTGTTGCGGTAATGTTGCA 1080
Db 1021 ACCGATATCTACGTCATGATGATTCCTCGTTTCTGTTTATTTAGTTGCGGTAATGTTGCA 1080
Qy 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGGTGTTGTCGATGAAGAAATTTAATCATTTGGGG 1140
Db 1081 TTTTCTAAAGAAATGGCTAAATAAAGTAGGTGTTGTCGATGAAGAAATTTAATCATTTGGGG 1140
Qy 1141 GCGCAAGATGTAGAAATTTGGTTACAGATATTTGCCAAAGGCTGTTTTCAGAGTAATT 1200
Db 1141 GCGCAAGATGTAGAAATTTGGTTACAGATATTTGCCAAAGGCTGTTTTCAGAGTAATT 1200
Qy 1201 GACGGCGAATGGCATCCATCAAGAACACCTGTTAAAGAAAAATGAACAGACCGGAA 1260
Db 1201 GACGGCGAATGGCATCCATCAAGAACACCTGTTAAAGAAAAATGAACAGACCGGAA 1260
Qy 1261 GCTGGTAAAGATTTACGCTTAAATTTGTGAAGAAAAGGTACCTTACATCTATAGAAG 1320
Db 1261 GCTGGTAAAGATTTACGCTTAAATTTGTGAAGAAAAGGTACCTTACATCTATAGAAG 1320
Qy 1321 CTTTACCATAGAGATTACATATTCATAGAACTTCTAGTTTCTATTTATATCCCC 1380
Db 1321 CTTTACCATAGAGATTACATATTCATAGAACTTCTAGTTTCTATTTATATCCCC 1380
Qy 1381 GCTTATAACTGCAAAATTTATATCAAGATGTGTAGATAGTGTCTTTAATCAAACTGTT 1440
Db 1381 GCTTATAACTGCAAAATTTATTCAGAGATGTGTAGATAGTGTCTTTAATCAAACTGTT 1440
Qy 1441 GTCGATCTCGAGTTTGTATTTGTAACGATGGTTCAACAGATATATACCTTAGAGTGATC 1500
Db 1441 GTCGATCTCGAGTTTGTATTTGTAACGATGGTTCAACAGATATATACCTTAGAGTGATC 1500
Qy 1501 AATAGCTTTATGTTAATTAATCTAGGTCAGCATATGTTCTAACCACAAATGGCGGAATA 1560
Db 1501 AATAGCTTTATGTTAATTAATCTAGGTCAGCATATGTTCTAACCACAAATGGCGGAATA 1560

Qy 1561 GCCTCAGCATCAAAATGCAGCCGTTCTTTTGTCTAAAGGTTTATTACATTTGGCAGTTAGAT 1620
Db 1561 GCCTCAGCATCAAAATGCAGCCGTTCTTTTGTCTAAAGGTTTATTACATTTGGCAGTTAGAT 1620
Qy 1621 TCAGATGATTATCTTGAGCCTGATGCAAGTTGAACCTGTGTTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATTATCTTGAGCCTGATGCAAGTTGAACCTGTGTTTAAAGAAATTTTAAAGAT 1680
Qy 1681 AAAACGCTAGCTTGTGTTTATACCACTTAATAGAAACCTCAATCCGGATGGTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTTGTGTTTATACCACTTAATAGAAACCTCAATCCGGATGGTAGCTTAATC 1740
Qy 1741 GCTAATGGTTTACAATTTGGCCAGAAATTTTTCACGAGAAAAAATCACAACGGCTATGATTGCT 1800
Db 1741 GCTAATGGTTTACAATTTGGCCAGAAATTTTTCACGAGAAAAAATCACAACGGCTATGATTGCT 1800
Qy 1801 CACATTTTGAATGTTTACGATTTAGAGCTTGGCATTTAAGGATTAACGGATGATTTAAGCAAAAT 1860
Db 1801 CACATTTTGAATGTTTACGATTTAGAGCTTGGCATTTAAGGATTAACGGATGATTTAAGCAAAAT 1860
Qy 1861 ATTGAAACCGCTGGGATTTATGACATGTTCTTAAACTCAGTGAAGTTGGAAAAATTTAAA 1920
Db 1861 ATTGAAACCGCTGGGATTTATGACATGTTCTTAAACTCAGTGAAGTTGGAAAAATTTAAA 1920
Qy 1921 CATCTTAAATAAATCTCTATAACCGGCTATTACATGGTGATACACATCCATTAAAGAAA 1980
Db 1921 CATCTTAAATAAATCTCTATAACCGGCTATTACATGGTGATACACATCCATTAAAGAAA 1980
Qy 1981 CTGCGGATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
Db 1981 CTGCGGATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
Qy 2041 ATCAATTTATTAATATATGACAAATTTGATGATTTAGATGAAAGTAGAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATATATGACAAATTTGATGATTTAGATGAAAGTAGAAGTATATCTTC 2100
Qy 2101 AATAAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTTAAACTCATTCAA 2160
Db 2101 AATAAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTTAAACTCATTCAA 2160
Qy 2161 AATAAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Db 2161 AATAAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Qy 2221 AAAAAACTTAAACAATTTATTTGAATATAATAAAAAATATTTCTGTTTATTTCTACATGTT 2280
Db 2221 AAAAAACTTAAACAATTTATTTGAATATAATAAAAAATATTTCTGTTTATTTCTACATGTT 2280
Qy 2281 GATAGAAATCATCTTACACGAGACATCAAAAAGAAATATTTGGCTTCTATCATPAAGCAC 2340
Db 2281 GATAGAAATCATCTTACACGAGACATCAAAAAGAAATATTTGGCTTCTATCATPAAGCAC 2340
Qy 2341 CAAGTGAATATTTTACTAAATAATGACATCTCATATTACACAGTAATAGACTAATAAAA 2400
Db 2341 CAAGTGAATATTTTACTAAATAATGACATCTCATATTACACAGTAATAGACTAATAAAA 2400
Qy 2401 ACTGAGGCACATTTAAGTAATTTAATAAATTAAGTCAAGTTAAATCTAAATTTGTGAATAC 2460
Db 2401 ACTGAGGCACATTTAAGTAATTTAATAAATTAAGTCAAGTTAAATCTAAATTTGTGAATAC 2460
Qy 2461 ATCAATTTTGTATATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Db 2461 ATCAATTTTGTATATCATGACAGCCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Qy 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTAAACACATGATTGGATCGAGAAAAATCAAT 2580
Db 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTAAACACATGATTGGATCGAGAAAAATCAAT 2580
Qy 2581 GCGCATCCACCATTTTAAAAAGCTGATTTAAACCTATTTTAAATGACAAATGACATTAAGAAGT 2640
Db 2581 GCGCATCCACCATTTTAAAAAGCTGATTTAAACCTATTTTAAATGACAAATGACATTAAGAAGT 2640

KW Hyaluronate synthase; PmHAS; hyaluronan; hyaluronic acid; HA; cosmetic;
KW drug delivery; angiogenesis; wound healing; capsule synthesis;
KW fowl cholera; shipping fever; ss.

OS Pasturella multocida.

XX Key Location/Qualifiers
XX CDS 19..2937
XX FT /*tag= a
XX FT /product= PmHAS
XX FT /note= "Hyaluronate synthase"

XX W09951265-A1.

XX PD 14-OCT-1999.

XX PF 01-APR-1999; 99WO-US07289.

XX PR 02-APR-1998; 98US-0080414.

XX PR 26-OCT-1998; 98US-0178851.

XX PA (OKLA) UNIV OKLAHOMA.

XX PI Deangellis P;

XX WPI; 2000-013032/01.

XX DR P-PSDB; AAY43099.

XX New isolated hyaluronate synthase nucleic acids, used for the
PT production of hyaluronic acid, for developing antibiotics and vaccines
PT and for diagnostic applications -

XX Claim 3; Page 113-114; 121pp; English.

XX This is the Pasturella multocida hyaluronate synthase (PmHAS) nucleic
CC acid sequence. Hyaluronic acid (HA) or hyaluronan, is a polysaccharide
CC that serves both structural and recognition roles in higher animals.
CC Bacteria produce extracellular capsules of HA which mimic their host HA
CC and aid escape from a host immune response. The invention includes a
CC vector containing the PmHAS nucleotide sequence which can be used to
CC express PmHAS in a foreign host. The HS nucleic acids can be used for the
CC production of HA. Also, specific changes to the HS coding sequence can
CC result in the production of HA having a modified size distribution or
CC structural configuration and functional properties. The HA products can
CC be used in e.g. drug delivery, angiogenesis and wound healing,
CC stabilisation of recombinant proteins and in cosmetics. The HS nucleic
CC acids can also be used to develop agents to block capsule synthesis by
CC pathogens and act as antibiotics. The avirulent P. multocida strains can
CC be used as vaccines for fowl cholera or shipping fever.

XX SQ Sequence 2937 BP; 1077 A; 482 C; 509 G; 869 T; 0 other;

Query Match 76.68; Score 2283.4; DB 21; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;

Qy 43 ATTTTAAAGGAAAGAAATGAATACATTATCACAGCAATAAAGACATATAACAGCAAT 102

Db 1 ATTTTAAAGGACAGAAATGAATACATTATCACAGCAATAAAGACATATAACAGCAAT 60

Qy 103 GACTATGAATTAGCAGCTCAAAATTTTGGAGAGTCTGCTGAAACCTACGGCGGAAAAATC 162

Db 61 GACTATCAATTAGCAGCTCAAAATTTTGGAGAGTCTGCTGAAACCTATGGAGGAAATTT 120

Qy 163 GTTGAATTCGAATATCAATGTAAAGAAAAAAGCTCTCGACCAATTC----- 209

Db 121 GTTGAATTTCAATTTACCAATGCCAAGAAAAAGCTCTCAGCACATCCCTCTGTTAATTC 180

Qy 210 -----TTATGTAAAGTGAAGATAAAAAACAGTGTTCGGATAGCTCATAGATATC 261

Db 181 GCACATCTTCTGAAATTAAGAGAAAAAGTCAATGTTTGGCATGTCGGTTAGATATT 240

Qy 262 GCAACACAGCTTACTTTTCCAAGCTAAAAAAATTAACCTCTATCCGAATCAGAAAAA 321

Db 241 GCAACACAACTTTACTTTCCAACTGTAATAAAATTAGTACTTCTGACTCGGAAAAAAC 300

Qy 322 AGTTTAAAAAATAATGGAATCTATCACTCGGAAAAAATCGAGAACGAGAAATCAGA 381

Db 301 ACGTTAAAAAATAATGGAATTTGCTCACTAGAGAAAAATCTGAAATCGGAGGTAAGA 360

Qy 382 AAGGTGGAACCTAGTACCCAAAGATTTTCCTAAAGATCTTGTCTTCTGCTCCATTCGCCAGAT 441

Db 361 GGGGTGCGCCTTGTACCAAAAGATTTTCCCAAAGATCTGTTTACGCGCTTACCTGAT 420

Qy 442 CATGTTAATGATTTTACATGGTACAAAAATCGAAAAAAGCTTAGGTATAAGACCTGTA 501

Db 421 CATGTTAATGATTTTACATGGTACAAAAAGCGAAAGAAAGACTTGGCATAAAACTGAA 480

Qy 502 AATAAGAAATATCGGTCTTCTATTTATTTCTACATTTTAACTCGTAGCCGTATTAGAT 561

Db 481 CATCAACATGTTGGTCTTTTCTATTTATCGTTACAACATTCATCGACGACGAAATTTATCG 540

Qy 562 ATAAGCTTAGCCTGTTTGGTCAATCAGAAAAAAGCTACCCATTTTGAAGTGTGTTGCA 621

Db 541 ATTACATTAGCCTGTTTGTAGTAACCAAAAAACACATTTACCGTTTGAAGTTATCGTGACA 600

Qy 622 GATGATGTAGTAAGAAAACTTACTTTACCATTTGTGCAAAAAATACGAAACAACTTTGAC 681

Db 601 GATGATGTAGTCAGGAAGATCTATCACCGATCATTCACGCAATATCGCAATATGAAATAAATGGAT 660

Qy 682 AATAAGTATGTAAAGAAAAAGATTTATGATATCATCAATTTGTGCGACTGACAGCAACAA 741

Db 661 ATTCGCTAGTCAGCAAAAAAGATAACGGTTTTTCAAGCCAGTCCGCTCGGAAATATGGGA 720

Qy 742 TTACGTACAGCAAAAGTATGATTTTGTCTCGATTTCTAGCTGCGATATGGCACCACAA 801

Db 721 TTACGCTTAGCAAAATATGACTTTATTGGCTTACTCGACTGATATGGCCCAATATCCA 780

Qy 802 TTATGGTTCATCTTATCTTACAGAACTATTAGAGAGCAATGATTTGTTTTAATTTGGA 861

Db 781 TTATGGTTCATCTTATCTTGTGAGAGCTATTAGAGAGATGATGTTTAAACAATCATTTGGT 840

Qy 862 CCTAGAAATATGTGGATCTCATATATATACCGCAGAACAAATCTCTTACGATCCATAT 921

Db 841 CCAAGAAAAATACGATCGATACACAACATATGACCCAAAAGACTTCTTAATAACGGGAGT 900

Qy 922 TTAATAGAACTACTACTCTGAAACCGCTACAAAATAACAATCTTTCGATTTACATCAAAAGGA 981

Db 901 TTGCTTGAATCATTACCAGAACTGAAACCAATAATAGTGTCCCGCAAAAGGGAGGA 960

Qy 982 AATATATCGTTGGATGGAGATTAGAACATTTCAAAAAACCGATTAATCTACCTCTATGT 1041

Db 961 ACAGTTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAAACAGAAATCTCCGCTTATCC 1020

Qy 1042 GATTCTCCGTTTCGTTATTTTGTTCGGGTAATGTTGCAATTTTCTAAAGATATGGCTAAAT 1101

Db 1021 GATTCCGCTTTCGTTTTCGCGGGTAAATGTTGCTTTCGCTTAAAAAATGGCTAAAT 1080

Qy 1102 AAAGTAGGTGTTGATCAAGAAATTTAATCATTTGGGGGGCGGAAGATGTAGAAATTTGGT 1161

Db 1081 AAATCCGGTTTCTTGTATGAGGAATTTAATCACTGGGGTGGAGAGATGTGGAATTTGGA 1140

Qy 1162 TCAGATTAATTTGCCAAAGCTGTTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCAT 1221

Db 1141 TATCGCTTATTCGTTTACGGTAGTTTCTTTTAAAACTATTGATGGCATTTATGGCTACCAT 1200

Qy 1222 CAAGAACACCTGGTAAAGAAAAATGAACAGACGCAAGCTGTAAGAGTATTACGCTT 1281

Db 1201 CAAGACCCAGGTAAAGAAAAATGAACCGCATCGTGAAGCGGGAAGAAATATTACGCTC 1260

Qy 1282 AAAATTTGCAAGAAAAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAAAGATTCA 1341

Db 1261 GATATTACAGAGAAAAAGGTCCCTTATATCTATAGAAAACTTTTACCAATAGAAAGATTTC 1320

Qy 1342 CATATTTCATAGAAATACCTTTTAGTTTCTATTATATATCCCGCTTATAACTGTGCAAAATAT 1401

Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATATCCAGCTTATAAAGTGTGCAAACTAT 1380
QY 1402 ATTCAAGATGTAGATAGTCTCTTATCAAACTGTGTGCACTCGAGCTTGTATTT 1461
Db 1381 ATTCAAGCTTGGCTAGATAGTCACTGAATCAGACTGTGTGATCTCGAGTGTGTATTT 1440
QY 1462 TGTACAGGATGTTCAACAGATATACCTTAGAAGTGATCAATAAGCTTTATGGTAAATAT 1521
Db 1441 TGTACAGATGTTCAACAGATATACCTTAGAAGTGATCAATAAGCTTTATGGTAAATAT 1500
QY 1522 CCTAGGTTACGATCATGTCTTAAACCAATAGCGGGAATAGCCTCAGCATCAAAATGAGCC 1581
Db 1501 CCTAGGTTACGATCATGTCTTAAACCAATAGCGGGAATAGCCTCAGCATCAAAATGAGCC 1560
QY 1582 GTTCTTTTGGCTTAAAGGTTATTTACATTTGGCAGTTAGATTCAGATGATTTATCTTGAGCC 1641
Db 1561 GTTCTTTTGGCTTAAAGGTTATTTACATTTGGCAGTTAGATTCAGATGATTTATCTTGAGCC 1620
QY 1642 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTAT 1701
Db 1621 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTAT 1680
QY 1702 ACCACTAATAGAACTCAATCGGATGTAGCTTAATCGCTAATGGTTACAAATTTGGCCA 1761
Db 1681 ACCACTAATAGAACTCAATCGGATGTAGCTTAATCGCTAATGGTTACAAATTTGGCCA 1740
QY 1762 GAATTTTTCAGGAAAGAACTCAACGCTGTATGATTCCTCACCATTGTTAGAACTGTTTACG 1821
Db 1741 GAATTTTTCAGGAAAGAACTCAACGCTGTATGATTCCTCACCATTGTTAGAACTGTTTACG 1800
QY 1822 ATTAGAGCTTGGCATTTAACGGATGATTAAACGAAATATTGAAACGCGCTGGATTTAT 1881
Db 1801 ATTAGAGCTTGGCATTTAACGGATGATTAAACGAAATATTGAAACGCGCTGGATTTAT 1860
QY 1882 GACATGTTCTTAACTCAGTGAAGTTGGAATTTTAAACATCTTAAATAAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTCAGTGAAGTTGGAATTTTAAACATCTTAAATAAATCTGCTAT 1920
QY 1942 AACCGGTATTACATGGTGATACACATCCATTAAAGAACTCGGCATTCAAAAGAAAAAC 2001
Db 1921 AACCGGTATTACATGGTGATACACATCCATTAAAGAACTCGGCATTCAAAAGAAAAAC 1980
QY 2002 CATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGCAATCAATTTATTAATATTATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGCAATCAATTTATTAATATTATGAC 2040
QY 2062 AAATTTGATGATTAGATGAAGTGAAGTATCTTCAATAAAACGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTAGATGAAGTGAAGTATCTTCAATAAAACGCTGAATATCAA 2100
QY 2122 GAAGAAATGGATGTTTAAAGATCTTAAACTCATTTCAAAATAAAGATGCCAAATCGCA 2181
Db 2101 GAAGAGATTGATATCTTAAAGATATTAATAATCATCCAGAATAAAGATGCCAAATCGCA 2160
QY 2182 GTCAGTATTTTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACAAATATTATT 2241
Db 2161 GTCAGTATTTTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACAAATATTATT 2220
QY 2242 GAATATAATAAATAATATTTCGTTATTATTCTACATCTTGAAGAATCATCTTACACCA 2301
Db 2221 GAATATAATAAATAATATTTCGTTATTATTCTACATCTTGAAGAATCATCTTACACCA 2280
QY 2302 GACATCAAAAAGAAATATTGGCTTTCTATCATAGAACCAAGTGAATATTTTACTAAAT 2361
Db 2281 GATATAAAAAGAAATATTAGCTTCTATCATAAACATCAAGTGAATATTTTACTAAAT 2340
QY 2362 AATGACATCTCATTTACAGAGTAATAGACTAATAAAACCTGAGGCACATTTTAAGTAAT 2421
Db 2341 AATGATATCTCATTTACAGAGTAATAGACTAATAAAACCTGAGGCACATTTTAAGTAAT 2400
QY 2422 ATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCATTTTGTATAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCATTTTGTATAATCATGAC 2460

QY 2482 AGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATATGATGCGCATGAAT 2541
Db 2461 AGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGCGCATGAAT 2520
QY 2542 TTCTCAGCATTAACACATGATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAG 2580
QY 2602 CTGATTAAAAACCTATTTTAAATGACAATGACTTAAAGAGTATGAATGTGAAGGGGCATCA 2661
Db 2581 CTCATTAAAAACCTATTTTAAATGACAATGACTTAAAGAGTATGAATGTGAAGGGGCATCA 2640
QY 2662 CRAAGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTATTAAGAAAGTC 2721
Db 2641 CRAAGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTATTAAGAAAGTC 2700
QY 2722 ATCAGATCTGCGCAATCAATTTGATAGTGTGCCAGATATACACTGAGGATATTTGGTTC 2781
Db 2701 ATCAGATCTGCGCAATCAATTTGATAGTGTGCCAGATATACACTGAGGATATTTGGTTC 2760
QY 2782 CAATTTGCACTTTTAACTCTTAGAAAAAGAAACCGGCCATGATTTTAAATAAAACATCGACC 2841
Db 2761 CAATTTGCACTTTTAACTCTTAGAAAAAGAAACCGGCCATGATTTTAAATAAAACATCGACC 2820
QY 2842 CTGACTTTATATGCTTGGGAAACGAAATTAACAATGGACAATGAACAAATTTCAAAGTCCA 2901
Db 2821 CTGACTTTATATGCTTGGGAAACGAAATTTACAATGGACAATGAACAAATTTCAAAGTCCA 2880
QY 2902 AAAAAAGCGCAAAATATCCCGTTTAAACAAGTTCAATTTTAAATAGTAAACGCTATAA 2958
Db 2881 AAAAAAGCGCAAAATATACCTGTTAAACAAGTTCAATTTTAAATAGTAAACGCTATAA 2937

RESULT 6
AAA27448
ID AAA27448 standard; cdna; 2112 BP.

XX AAA27448;

DT 17-AUG-2000 (first entry)

DE P.multocida hyaluronic acid synthase-D coding sequence.

XX Hyaluronic acid synthase-D; HAS-D; enzyme; ophthalmic surgery;
KW cataract; arthritis; ulcer; tissue abrasion; viscoelastic replacement;
KW hyaluronic acid production; bioadhesive; ss.

OS Pasteurella multocida.

XX Key Location/Qualifiers
FH CDS 1..2112
FT /*tag= a
FT /product= "PmHAS-D"
FT /trans_except= (pos:1348..1353,aa:Asn)

XX WO200027437-A2.

XX 18-MAY-2000.

XX 10-NOV-1999; 99WO-US26501.

XX 11-NOV-1998; 98US-0107929.

XX 01-APR-1999; 99US-0283402.

XX (OKLA) UNIV OKLAHOMA STATE.

XX DeAngelis PL;

XX WPI; 2000-376319/32.

XX P-PSDB; AAY96212.

XX Novel method for the enzymatic transfer of sugar molecules to an

FT CDS /note= "encodes AAY54065"
 3390..4121
 /*tag= b
 /product= "EPS2"
 /note= "encodes AAY54066"
 4130..4822
 /*tag= c
 /product= "EPS3"
 /note= "encodes AAY54067"
 4832..5530
 /*tag= d
 /product= "EPS4"
 /note= "encodes AAY54068"
 5629..6996
 /*tag= e
 /product= "EPS5"
 /note= "encodes AAY54069"
 8633..10681
 /*tag= f
 /product= "EPS6"
 /note= "encodes AAY54070: contains 1 stop codon at
 nucleotides 9410-9412"
 10703..11671
 /*tag= g
 /product= "EPS7"
 /note= "encodes AAY54071"
 12838..13788
 /*tag= h
 /product= "EPS8"
 /note= "encodes AAY54072"
 14138..15553
 /*tag= i
 /product= "EPS9"
 /note= "encodes AAY54073"
 16919..18016
 /*tag= j
 /product= "EPS10"
 /note= "encodes AAY54074"
 W09962316-A2.
 09-DEC-1999.
 22-APR-1999; 99WO-EP02841.
 22-APR-1998; 98EP-0201310.
 22-APR-1998; 98EP-0201311.
 22-APR-1998; 98EP-0201312.
 (NEST) SOC PROD NESTLE SA.
 Stinglele F, Germond JE, Lamothe G;
 WPT; 2000-097267/08.
 P-PSDB; AAY54065, AAY54066, AAY54067, AAY54068, AAY54069, AAY54070,
 AAY54071, AAY54072, AAY54073, AAY54074.
 New recombinant enzymes for synthesis of exopolysaccharides,
 particularly in lactic acid bacteria, for improving properties of
 fermented milk products -
 Claim 6; Page 77-96; 162pp; French.
 The present sequence encodes enzymes involved in the biosynthesis of
 exopolysaccharides (EPS). These enzymes are designated EPS1-EPS10. and
 are encoded by open reading frames eps1-eps10. The enzymes are isolated
 from *Streptococcus thermophilus* strain Sfi39. The proteins are used
 in a method for the synthesis of EPS, which includes at least one step
 of forming a bond (alpha or beta-isomer) between C-1 (carrying the
 reducing aldehyde function, of an activated D-galactose pyranose), and
 a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis
 of EPS occurs with, in each step, addition of a new sugar unit, through
 its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar

CC unit, present at the end of a chain of sugar residues bonded to the
 CC primer. EPS1 to EPS4 are used to elongate the EPS chain, and to regulate
 CC EPS production, EPS5 creates new bonds between saccharides, EPS6 is used
 CC in the biosynthesis of EPS, EPS7 and EPS8 are beta-glycosyltransferases,
 CC EPS9 is a transporter of repetitive units, and EPS10 catalyses the
 CC conversion of a pyranose form of a beta-D-galactose to the furanose
 CC form. The EPS enzyme are used to improve properties of foods,
 CC particularly fermented milk products such as yoghurt and cheese,
 CC e.g. their organoleptic properties and flavour stability.
 XX
 SQ Sequence 19966 BP; 6405 A; 3134 C; 3732 G; 6694 T; 1 other;

Query Match 2.7%; Score 80; DB 21; Length 19966;
 Best Local Similarity 52.0%; Pred. No. 4.1e-06;
 Matches 179; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

OY 1307 ACATCTATAGAAAGCTTTTACCAATAGAGATTACACATATTCATAGAAATACCTTTAGCTTT 1366
 DB 12794 ACTGTTAGAAATATGATTTTATATAATTTAGGAGTAGATAAAGAGATGAATCCATTATAT 12853
 OY 1367 CTATTTATATCCCGCTTATACTGTGCAAAATATATATCAAAAGATGTGTAGATGCTC 1426
 DB 12854 CAATTATTGTTCCAACATATAATGTTGAAAAATATATTAGACATGATTGAATCAATCT 12913
 OY 1427 TTAATCAAACTGTTGCGATCTCGAGGTTTGTATTGTAACGATGCTTCACAGATATA 1486
 DB 12914 TAGCTCAACATATCGCAATATTGAAGTCATTATAGTATGATGGTAGCAGATCAGT 12973
 OY 1487 CCTAGAGATGATCAATAAGCTTTATGTTAATATCTAGGTTACGATCATGTCTAAAC 1546
 DB 12974 CGTAGCAGTAATTTCCGATTTAATCTGTACTCATCATATATTAAAGTATATCAACCAA 13033
 OY 1547 CAAATGCGGAATAGCCTCAGCATCAAAATGACGCGGTTCTTTTGTCTAAAGGTTATTACA 1606
 DB 13034 AAACCAAGGATTATATAGTAGTCTGAAACACTGTTATGATCGCGCACTGCTAAATATA 13093
 OY 1607 TTGGCGAGTTAGATTTCAGATGATTATCTTGAGCTCGATGTCAGTT 1650
 DB 13094 TAGCTTTGTAGATGCGAGTACAAATAATTAAGCCAGACTTTGTT 13137

RESULT 11

AAY30355
 ID AAY30355 standard; DNA; 19966 BP.

XX AC AAY30355;

XX DT 11-FEB-2000 (first entry)

XX DE Nucleotide sequence of the eps operon of *S. thermophilus* Sfi39.

XX KW eps operon; *Streptococcus thermophilus* Sfi39; enzyme; eps1; eps2;
 KW eps3; eps4; eps5; eps6; eps7; eps8; eps9; eps10;
 KW exopolysaccharide biosynthesis; EPS; intersugar bond; antitumour;
 KW probiotic; foodstuff; organoleptic quality; flavour;
 KW lactic acid bacteria; acidified milk product; yoghurt; cheese; ss.

XX OS *Streptococcus thermophilus*.

XX FH Key Location/Qualifiers

FT CDS 2703..3389
 FT /*tag= a
 FT /product= "eps1"
 FT /note= "encodes AAY43767"
 FT 3390..4121
 FT /*tag= b
 FT /product= "eps2"
 FT /note= "encodes AAY43768"
 FT 4130..4822
 FT /*tag= c
 FT /product= "eps3"
 FT /note= "encodes AAY43769"
 FT 4832..5530
 FT CDS


```
FT FT /*tag= d
FT FT /product= "eps4"
FT FT /note= "encodes AAY43770"
FT CDS 5629..6996
FT FT /*tag= e
FT FT /product= "eps5"
FT FT /note= "encodes AAY43771"
FT CDS 8633..10681
FT FT /*tag= f
FT FT /product= "eps6; this CDS, minus the termination
FT FT . codon, is specifically claimed (claim 6)"
FT FT /note= "encodes AAY43772; contains a stop codon at
FT FT nucleotides 9410-9412"
FT CDS 10703..11671
FT FT /*tag= g
FT FT /product= "eps7"
FT FT /note= "encodes AAY43773; this CDS, minus the
FT FT termination codon, is specifically claimed
FT FT (claim 6)"
FT CDS 12838..13788
FT FT /*tag= h
FT FT /product= "eps8"
FT FT /note= "encodes AAY43774; this CDS, minus the
FT FT termination codon, is specifically claimed
FT FT (claim 6)"
FT CDS 14138..15553
FT FT /*tag= i
FT FT /product= "eps9"
FT FT /note= "encodes AAY43775; this CDS is specifically
FT FT claimed (claim 6)"
FT CDS 16919..18016
FT FT /*tag= j
FT FT /product= "eps10"
FT FT /note= "encodes AAY43776; this CDS, minus the
FT FT termination codon, is specifically claimed
FT FT (claim 6)"
FT FT
FT XX W09954475-A2.
FT XX
FT XX
FT XX 28-OCT-1999.
FT XX
FT XX 22-APR-1999; 99WO-EP03011.
FT XX
FT XX 22-APR-1998; 98EP-0201310.
FT PR 22-APR-1998; 98EP-0201311.
FT PR 22-APR-1998; 98EP-0201312.
FT XX
FT PA (NEST ) SOC PROD NESTLE SA.
FT XX
FT XX Stingele F, Germond JE, Lamothe G;
FT XX
FT XX WPI: 2000-013255/01.
FT DR P-PSDB; AAY43767, AAY43768, AAY43769, AAY43770, AAY43771, AAY43772,
FT DR AAY43773, AAY43774, AAY43775, AAY43776.
FT XX
FT XX New recombinant enzymes for biosynthesis of exopolysaccharides having
FT PT e.g. antitumor or probiotic properties or useful in fermented milk
FT PT products
FT XX
FT PS Claim 6; Page 77-97; 163pp: French.
FT XX
FT CC The present sequence represents the eps operon of Streptococcus
FT CC thermophilus Sfi139. The operon contains 10 open reading frames, and
FT CC encodes enzymes (eps1, eps2, eps3, eps4, eps5, eps6, eps7, eps8, eps9
FT CC and eps10) that are involved in the biosynthesis of exopolysaccharides
FT CC (EPS). The enzymes catalyse the formation of specific intersugar bonds.
FT CC The enzymes catalyse a process that includes at least one step of
FT CC forming a bond (in alpha or beta anomeric form) between C1, carrying
FT CC the reducing aldehyde group of an activated D-Galp (galactose in
FT CC pyranose form), and a phosphate group on a lipophilic or proteinaceous
FT CC primer. The enzymes are used to produce EPS that have antitumor or
FT CC probiotic properties or are used in foodstuffs to improve organoleptic
FT CC qualities and flavour. When expressed by lactic acid bacteria, EPS impart
```

```
CC a free-flowing character and/or a smooth, creamy texture to acidified
CC milk products (yoghurt or cheese).
XX
SQ Sequence 19966 BP; 6405 A; 3133 C; 3732 G; 6695 T; 1 other;
Query Match 2.7%; Score 80; DB 21; Length 19966;
Best Local Similarity 52.0%; Pred. No. 4.1e-06;
Matches 179; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
QY 1307 ACATCTATAGAAGCTTTTACCAATAGAAGATTCACATATTCATAGAATACCTTTAGTTT 1366
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12794 ACTGTTAGAATATGATTTTATATAATTAGGAGTAGAATAAAGAGATCAATCCATTAAT 12853
QY 1367 CTATTTATATCCCGCTTATAAAGTGTGCAAAATTATATTCAAAGCATGTGTAGATGCTC 1426
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12854 CAATTATGTTCCACATATATATGTTGAAAATATATATTAGGACATGATTTGAATCAATCT 12913
QY 1427 TTAATCAAACTGTTGTCGATCTCGAGGTTTGTATTTTGAACGATGGTTCAACAGATAATA 1486
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12914 TAGCTCAAAACATATCGCAATATTGAAGTCATTTATAGTAAATGTTGTAGCACAGATCAGT 12973
QY 1487 CCTTGAAGTGTATCAATAGCTTTATGTTAATTAATCCTTAGGGTACGCATCATGTCTAAAC 1546
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12974 CGCTAGCAGTAATTTCCGATTTAATCTGTAGTCATCATATAATTAAGGTAATCAACCAA 13033
QY 1547 CAAATGGCGGAATAGCCTCAGCATCAAATGCAGCCGTTTCTTTTGTCTAAAGGTTATTACA 1606
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13034 AAACCAAGATTTATCAGTAGCTCGAAACACTGGTATTGATGCGGCAACTGGTAATATA 13093
QY 1607 TTGGCAGTTAGATTCAGATGATTATCTTTGAGCCTGATGCAGTT 1650
FT ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13094 TAGCTTTTGTAGATGCAGATGACAAAATTAAGCCAGACTTTTGT 13137
RESULT 12
AAY43720/c
ID AAY43720 standard; DNA; 11474 BP.
XX
AC AAY43720;
XX
DT 22-DEC-2000 (first entry)
XX
DE LPS core biosynthesis locus.
XX
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside; ds.
XX
OS Campylobacter jejuni OH4384.
XX
FH Key Location/Qualifiers
FT CDS 350..1237
FT FT /*tag= a
FT FT /product= acyltransferase
FT FT /note= "Open reading frame 2b"
FT CDS 1234..2490
FT FT /*tag= b
FT FT /product= glycosyl transferase
FT FT /note= "Open reading frame 3a"
FT CDS 2786..3955
FT FT /*tag= c
FT FT /product= Glycosyl transferase
FT FT /note= "Open reading frame 4b"
FT CDS 4025..5068
FT FT /*tag= d
FT FT /product= Beta 1,4-GalNAC transferase
FT FT /note= "Open reading frame 5a"
FT CDS 5054..5959
FT FT /*tag= e
FT FT /product= Beta 1,2- galactosyltransferase
FT FT /note= "Open reading frame 6a"
```


QY 2532 CGGCATGATTTCTCAGCATTAACACATGATTTGGATCGAGAAAAATCAATGCGCATCCACC 2591
DB 16467 TAGCACCACAACTAACCAATAACGAAATCCCGTCTCTACTATAAACACACACAAAAA 16408
QY 2592 ATTAAAAAGCTGATTAAACCTATTTTAATGACATGACTTAAAGAGTATGAATGTGAA 2651
DB 16407 TTAACGAACATANTAACACACACCTATATCCCACTACTCAAAAACTAAAAATAAAT 16348
QY 2652 AGGGGCATCACAGGATGTTTATGAAGTATGCGGTACCGCATGAGCTTCTGACGATTAT 2711
DB 16347 TACTTAAACCCAAAAATAAAAAATTACTATAAACCTTAAATCTTACCATTTACATCCAAC 16288
QY 2712 TAAAGAGTCATCACATCTCGCAATCAATTTGATAGTGTGCGAGATATAAACAACCTCAGGA 2771
DB 16287 TAAACCAAAAAACGAACTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 16228
QY 2772 TATTTGGTTCCTAATTTGCACTTTTATCTTAGAAAAAGAAACCGGCCATGTTATTAATA 2831
DB 16227 ATATAATCTCTATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16168
QY 2832 ACATCGACCTGACTTATATGCTTTGGGAACGAAATTTACATGACACAAATGAACAAAT 2891
DB 16167 AACCTTTTACCAATAAAAAACCAAAAAACAAAAAATAAATAAATAAATAAATAAATAAATAA 16108
QY 2892 TCAAGTGCAAAAAGGCGAAATATCCCGTTTAAACAAGTTTCATTTATTAATAGTATAAC 2951
DB 16107 TATCCCTATTAAAAATCAAACTATCTCATATATTTCTTCAACATACTATAAATA 16048
QY 2952 GCTATAAAACATTTGCATTTTATTAATAA 2979
DB 16047 CAAAAATATATCCCAATTTTATAACA 16020

RESULT 15
AAS63332/c
ID AAS63332 standard; DNA; 17527 BP.

XX
AC AAS63332;
XX
DT 29-JAN-2002 (first entry)

XX Chemically pretreated metabolism associated gene #27.

XX Human; cytostatic; anti-tumour; metabolism; metabolic disease; liver;
XX solid tumour; cancer; cytosine methylation; epigenetic; eye; kidney;
XX single nucleotide polymorphism detection; SNP; stool; urine; lung;
XX cerebral-spinal fluid; intestine; brain; heart; prostate; breast;
XX DUSP2; EPHX2; QDPR; SGSH; SHMT2; SLC7A2; SLC7A4; TYMS; ds.

XX Homo sapiens.

XX W0200176451-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EF04016.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-010834/01.

XX New nucleic acid, useful for diagnosis and therapy of metabolic

XX disease, solid tumour and cancers, comprises segment of chemically

XX modified genomic sequences of genes associated with metabolism

XX Claim 1; Page 79-84; 143pp; English.

XX

CC The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases of a segment of the chemically pretreated DNA of genes
CC associated with metabolism such as DUSP2 (NM_004418), EPHX2 (NM_001979),
CC QDPR (NM_000320), SGSH (NM_000199), SHMT2 (NM_005412), SLC7A2
CC (NM_003046), SLC7A4 (NM_004173) and TYMS (NM_001071) (all
CC undefined). (I) are useful for diagnosis and therapy of metabolic
CC disease, solid tumours and cancers; as primer oligonucleotides for the
CC amplification of DNA sequences, for detecting the cytosine methylation
CC state and/or single nucleotide polymorphisms (SNPs) in a chemically
CC treated DNA of genes associated with metabolism. An array of (I) is
CC useful for ascertaining genetic and/or epigenetic parameters for the
CC diagnosis and/or therapy of existing diseases or the predisposition to
CC specific diseases by analysing cytosine methylations. The method involves
CC chemically treating genomic DNA sample by a solution of bisulphite,
CC hydrogen sulphite or disulphite such that cytosine bases which are
CC un methylated at the 5th position are converted to uracil or another base
CC which is dissimilar to cytosine in terms of hybridisation behaviour and
CC amplifying fragments of the chemically pretreated genomic DNA. The
CC genomic DNA is from cells or cellular components which contain DNA,
CC sources of DNA comprising, for e.g. cell lines, biopsies, blood, sputum,
CC stool, urine, cerebral-spinal fluid, tissue embedded in paraffin such as
CC tissue from eye, intestine, kidney, brain, heart, prostate, lung, breast
CC or liver, histologic object slides and their combinations. Genetic
CC parameters are mutations, in particular insertions, deletions, point
CC mutations, inversions and polymorphisms of genes associated with
CC metabolism and sequences further required for their regulation.
CC Epigenetic parameters are in particular cytosine methylations and
CC further chemical modifications of DNA bases of genes associated with
CC metabolism. Further epigenetic parameters include for e.g. the
CC acetylation of histones which correlates with DNA methylation.
CC AAS63306-AAS63373 represent chemically pretreated metabolism associated
CC genes, and related primers of the invention.

XX SQ Sequence 17527 BP; 4501 A; 296 C; 4094 G; 8636 T; 0 other;

Query Match 2.68; Score 77.2; DB 24; Length 17527;
Best Local Similarity 43.9%; Pred. No. 1.4e-05;
Matches 434; Conservative 0; Mismatches 543; Indels 11; Gaps 2;

QY 1992 AAGAAAAAACCATTTTGTGTAGTCAATCAGTCATTAATAAGACAAAGGCATCAATTA 2051

DB 16996 ATAAAAAACGGTATCTTAATACCTAAAAAATTAATAACCAACACGA 16937

QY 2052 TAATATGACAAATTTGATGATGATGAAAGTAGAAGTATATCTTCAATAAACCCG 2111

DB 16936 TAATACACCTTAAATCCCACTACTCAAAAAACAAAAATAATATTACTTAAACCC 16877

QY 2112 TGAATATCAGAAGAAATGGATATGTTAAAAAGATCTTAAACTCATTTCAAAATAAAGATGC 2171

DB 16876 AAAATTTAAAAATAACAATAAACCAAAATCACACCCCTATCTCCAACTTAACACAAAA 16817

QY 2172 CAAAAATCGCAGTCAGTATTTTCTATCCCAATACATTAACCGCTTAGTGAAAAAACTAAA 2231

DB 16816 CAAACCCCACTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 16757

QY 2232 CAATATTTTGAATATATAATAATAATAATATTCGTTATTTTCTACATGTTGATAAGATCA 2291

DB 16756 AAAAAACAACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 16701

QY 2292 TCTTACACCGACATCAAAAAAGAAATATTGGCTTCTTATCATATAAGCAGCAAGTGAATAT 2351

DB 16700 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 16641

QY 2352 TTTTACTAAATAATGACATCTCATATTTACACGAGTAAATAGACTTAATAAACCTGAGGCACA 2411

DB 16640 TTTCTAAATAATACGACTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 16588

QY 2412 TTTAAGTAAATAATAAATAAATTAAGTCAGTTAAATCTAAATTTGATGATCATTTTCA 2471

DB 16587 TTTTAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 16528

QY 2472 TAATCATGACACGCTTATCGTTTAAAAAATGACAGCTATGCTTTATATGAAAAAATAATGATGT 2531

```
Db 16527 TACCTATAATCCTAACACTTTTAAAAATCAAAACAAATAAATCACCTAAAAATCAAAAAAT 16468
Qy 2532 CGGCATGAATTTCTCAGCATTAACACATGATTGGATCCAGAAAAATCAATGGCATCCACC 2591
Db 16467 TACGACCAACTAACCAAAATACGAAATCCCGTCTCTACTAAAAACACACACAAAAAA 16408
Qy 2592 ATTTAAAAAGCTGATTAAAACTTTTAAATGACAATGACTTAAGAAGTATGAATGTGAA 2651
Db 16407 TTAACGAAACATAATAACACACACCTATAATCCCAACTACTCARAAAACTAAAAATAAAT 16348
Qy 2652 AGGGGCATCAAGGTATGTTATGAAGTATGCGCTACCGCATGAGCTCTGACCATTTAT 2711
Db 16347 TACTTAAACCCAAAAATAAAAAATTTACTTATAAACCTAAATCCTACCATTACACTCCAACC 16288
Qy 2712 TAAAGAAGTCATCACATCCTGCCAATCAATTGATAGTGTGCCAGAAATATAACACTGAGGA 2771
Db 16287 TAAACACAAAAACGAAACTCCTCAAAAAAATTAATTAATTTAATTTAAAAA 16228
Qy 2772 TATTTGGTTCCAATTTGCACCTTTTAACTTAGAAAAAGAAACCGGCCATGTATTTAATAA 2831
Db 16227 ATATAATCTCTATATAAAATAAAAATAAAAAATACAAAAATATATATTTATAAAATAA 16168
Qy 2832 AACATCGACCTGACTTATATGCCTTGGGAACGAAAAATTACAATGGACAAAATGAACAAAT 2891
Db 16167 AACCTTTACCAATAAAAAACCAAAAAACAAAAAACTATAAACGAAAAATTTATTACTA 16108
Qy 2892 TCAAAGTCAAAAAAGCGGAAATATCCCGTTAACAAAGTTCAATTATATAGTATTAAC 2951
Db 16107 TATCCCTATTAAAAATCAAAACTATCTCTCATATATTCCTTCAAAACATAACTATAAATTA 16048
Qy 2952 GCTATAAACACATTTGCATTTTATTAAAA 2979
Db 16047 CAAAAATATTATCCCAATTTTATAACA 16020
```

Search completed: January 3, 2003, 21:32:20
Job time : 737 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:18:18 ; Search time 117 Seconds
(without alignments)
7808.462 Million cell updates/sec

Title: US-09-842-484A-1
Perfect score: 2979
Sequence: 1 ttataactgattaaagaag.....acatttcattttattataaa 2979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2979	100.0	2979	4	US-09-437-277-4
2	1503.2	50.5	2112	4	US-09-437-277-2
3	74.6	2.5	14602	1	US-08-597-236-1
4	74.6	2.5	14602	1	US-08-746-682A-1
c 5	74.2	2.5	615	4	US-08-998-416-186
c 6	70.6	2.4	837	4	US-08-998-416-288
c 7	70.2	2.4	636	4	US-08-998-416-1137
8	67.2	2.3	32768	4	US-08-961-527-71
9	66.8	2.2	1956	4	US-08-559-896B-1
10	63.2	2.1	19124	2	US-08-487-826B-13
c 11	61	2.0	19390	4	US-08-961-527-86
12	60.6	2.0	2023	3	US-08-961-083-199
c 13	58.8	2.0	658	4	US-08-998-416-595
c 14	57.8	1.9	6243	2	US-09-056-075-1
c 15	57.6	1.9	811	3	US-08-961-083-197
c 16	55.6	1.9	7218	1	US-08-232-463-14
c 17	55	1.8	14187	4	US-09-453-702B-121
c 18	54.2	1.8	5852	1	US-07-867-106-2
c 19	53.8	1.8	663	4	US-08-998-416-191
20	53.6	1.8	8920	2	US-08-446-855A-1
21	53.6	1.8	8920	4	US-09-150-741-1
22	52.2	1.8	854	4	US-08-998-416-534
c 23	52.2	1.8	168575	4	US-09-426-290-1
c 24	51.6	1.7	724	4	US-08-998-416-683
c 25	51	1.7	711	4	US-08-998-416-786
26	51	1.7	3095	6	5231168-1
27	51	1.7	3480	1	US-08-920-812-11

28	51	1.7	3480	1	US-08-920-827-11	Sequence 11, Appl
29	51	1.7	3480	1	US-08-921-177-11	Sequence 11, Appl
30	51	1.7	3480	1	US-08-362-577C-11	Sequence 11, Appl
31	51	1.7	3480	2	US-08-920-828-11	Sequence 11, Appl
c 32	50.4	1.7	1422	1	US-08-319-704-5	Sequence 5, Appl
c 33	50.2	1.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
34	49.8	1.7	1431	4	US-09-316-083-2	Sequence 2, Appl
35	48.8	1.6	827	4	US-08-998-416-535	Sequence 535, App
36	48.8	1.6	5181	1	US-08-257-073-10	Sequence 10, Appl
37	48.6	1.6	660	1	US-07-991-867B-32	Sequence 32, Appl
38	48.6	1.6	660	1	US-08-107-755A-32	Sequence 32, Appl
39	48.6	1.6	660	2	US-08-544-332-32	Sequence 32, Appl
40	48.6	1.6	660	4	US-09-370-861A-32	Sequence 32, Appl
41	48.6	1.6	1511	1	US-07-991-867B-8	Sequence 8, Appl
42	48.6	1.6	1511	1	US-08-107-755A-8	Sequence 8, Appl
43	48.6	1.6	1511	2	US-08-544-332-8	Sequence 8, Appl
44	48.6	1.6	1511	4	US-09-370-861A-8	Sequence 8, Appl
45	48.6	1.6	4507	2	US-08-568-459A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-437-277-4
; Sequence 4, Application US/09437277
; Patent No. 6444447
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 5820.551
; CURRENT APPLICATION NUMBER: US/09/437,277
; CURRENT FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
; SEQ ID NO 4
; LENGTH: 2979
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-09-437-277-4

Query Match 100.0%; Score 2979; DB 4; Length 2979;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTATAA	CTGATTAAAGAGGTTAAACGATTCAAGCAAGGTTAAATTTTAAAGGAAGAAA	60
DB	1	TTATAA	CTGATTAAAGAGGTTAAACGATTCAAGCAAGGTTAAATTTTAAAGGAAGAAA	60
QY	61	ATGAAT	TACATTATCAACAGCAATAAAGCATATACAGCAATGACTATGATTTAGCACTC	120
DB	61	ATGAAT	TACATTATCAACAGCAATAAAGCATATACAGCAATGACTATGATTTAGCACTC	120
QY	121	AAATTT	TTGAGAAGTCTGCTGAAACCTACGGCGGAAAAATCGTTGAAATTTCCAAATATC	180
DB	121	AAATTT	TTGAGAAGTCTGCTGAAACCTACGGCGGAAAAATCGTTGAAATTTCCAAATATC	180
QY	181	AAATGT	AAAGAAACTCTCGACCAATTTCTATGTAAGTGAACATATAAAGAAACAGTGT	240
DB	181	AAATGT	AAAGAAACTCTCGACCAATTTCTATGTAAGTGAACATATAAAGAAACAGTGT	240
QY	241	TGCGAT	GATCTATTAGATATCGCAACACAGCTCTTACTTTCCAACTGTAATAAATAAAT	300
DB	241	TGCGAT	GATCTATTAGATATCGCAACACAGCTCTTACTTTCCAACTGTAATAAATAAAT	300
QY	301	CTATCG	GAATCAGAAAAAACAGTTTAAATAAATGGAATCTATCTACTGGGAAAAAA	360
DB	301	CTATCG	GAATCAGAAAAAACAGTTTAAATAAATGGAATCTATCTACTGGGAAAAAA	360
QY	361	TCGGAG	AGCCGAGAAATCAGAAAGTGGAACTAGTACCCAAAGATTTTCTTAAGATCTT	420
DB	361	TCGGAG	AGCCGAGAAATCAGAAAGTGGAACTAGTACCCAAAGATTTTCTTAAGATCTT	420

Db 2581 GCACATCCACCATTAAAAAGCTGATTAACAACTATTATTAATGACAATGACTTAAGAGT 2640
Qy 2641 ATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Qy 2701 CTGACGATTTAAAGAAGTCATCACATCTCTGCCAATCAATGTAGTGTGCCAGATAT 2760
Db 2701 CTGACGATTTAAAGAAGTCATCACATCTCTGCCAATCAATGTAGTGTGCCAGATAT 2760
Qy 2761 AACACTGAGGATTTTGGTCCCAATTTGACATTTTATCTAGAAAAGAAACCGGCAT 2820
Db 2761 AACACTGAGGATTTTGGTCCCAATTTGACATTTTATCTAGAAAAGAAACCGGCAT 2820
Qy 2821 GTATTTAAATAACATCGACCTGACATATATGCTTTGGGAACCAAAATTAACAATGGACA 2880
Db 2821 GTATTTAAATAACATCGACCTGACATATATGCTTTGGGAACCAAAATTAACAATGGACA 2880
Qy 2881 AATGAACAAATTCAAAGTGCACAAAGGGCGAAATATCCCGGTTAACAAAGTTCATTATT 2940
Db 2881 AATGAACAAATTCAAAGTGCACAAAGGGCGAAATATCCCGGTTAACAAAGTTCATTATT 2940
Qy 2941 AATAGTATACGCTATAAACAATTTGCATTTTATTAATA 2979
Db 2941 AATAGTATACGCTATAAACAATTTGCATTTTATTAATA 2979

RESULT 2
US-09-437-277-2
; Sequence 2, Application US/09437277
; Patent No. 644447
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 5820.551
; CURRENT APPLICATION NUMBER: US/09/437, 277
; CURRENT FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
; SEQ ID NO 2
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-09-437-277-2

Query Match 50.5%; Score 1503.2; DB 4; Length 2112;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 348; Indels 21; Gaps 1;

Qy 61 ATGAATACATTTATCACAGCAATTAAGAGCATATAACAGCAATGACTATGAATTAGCACTC 120
Db 1 ATGAATACATTTATCACAGCAATTAAGAGCATATAACAGCAATGACTATGAATTAGCACTC 60
Qy 121 AAATTTATTTGAGAGTCTGCTGAAACCTACGGCGGAAATAATCGTTGAATCCAAATATAC 180
Db 61 AAATTTATTTGAAAGTGGCGGAAATCTATGGACGGAAATTTGTTGAATTTCAAATTACC 120
Qy 181 AAATGTAAGAAAACCTCTGCACCAATTC-----TTATGTAAGT 219
Db 121 AAATGCAAGAAAACCTCTAGCAGACATCTCTGTTAATTCAGCACATCTTTTCTGTAAAT 180
Qy 220 GAAGATAAAAAACAGTGTTCGGATAGCTCAATAGATATCGCAACACAGCTCTTACTT 279
Db 181 AAAGAAGAAAAGTCAATGTTTGGCATAGTCCGTTAGATATGGCAACACTGTTACTT 240
Qy 280 TCCACGTAATAAAATTAACCTCTATCCGAATCAGAAAAACAGTTTAAAAAATAAATG 339
Db 241 TCCAACGTAATAAAATTAAGTACTTCTGACTCGGAAAAACACGTTAAAAAATAAATG 300
Qy 340 AATCTATCACTGGGAAAAATTCGGAACCGCAGAAATTCAGAAAGGTGGAACCTAGTACC 399
Db 301 AAATTTGCTCACTGAGAAGAAATCTGAAANTGCGGAGGTGAAGCGGCTGCCCTTTGTACCA 360

Qy 400 AAAGATTTTCTTAAGATCTTTGTTCTTGTCCATTCAGATGATCAATGTTTAAATGATTTTACA 459
Db 361 AAAGATTTTCTTAAGATCTTTGTTCTTGTCCATTCAGATGATCAATGTTTAAATGATTTTACA 420
Qy 460 TGGTACAAAATTCGAAAAAAGCTTAGGTATAAAGCCCTGTAAATGAAGAAATATCGGTCTTT 519
Db 421 TGGTACAAAATTCGAAAAAAGCTTAGGTATAAAGCCCTGTAAATGAAGAAATATCGGTCTTT 480
Qy 520 TCTATTATTATTTCCATACATTTAAATCGTAGCCGATTTTATAGATATAACGTTTACCCCTGTTTG 579
Db 481 TCTATTATCTGTACAAACATTTCAATCGCAGCAATTTTATCGATTTACATTAACCTGTTTAT 540
Qy 580 GTCATATCAAAAAACAACTACCCATTTTGAAGTCGTTTGTTCAGATGATGGTAGTAGGAA 639
Db 541 GTCATATCAAAAAACAACTACCCATTTTGAAGTCGTTTGTTCAGATGATGGTAGTAGGAA 600
Qy 640 AACTTACTTTACCATTTGTCAAAAATACGACAAAAAATTCGATATAAGTATGTAAGACAA 699
Db 601 GATCTATCAACGATCATTTCCGCAATATGAAAAATAAATTTGGATATTTCCGCTACGTCAGACAA 660
Qy 700 AAAGATTTATGGATATCAATTTGTGTCAGTTCAGAAACTTAGGTTTACGTACACGCAAGATAT 759
Db 661 AAAGATTAACGTTTTCAGCCAGTCCGCTCGGAATATGGGATTTAGCGTTACGCAAAATAT 720
Qy 760 GATTTTGTCTCGATTTCTAGACTCGATATGGCACCACACAAATTTATGGGTTTCATTTCTTAT 819
Db 721 GACTTTATTTGGCTTACTCGACTGTGATATGGCGCAAAATCCATTTATGGGTTTCATTTCTTAT 780
Qy 820 CTTACAGAACTATTAGAAGACAAATGATTTGTTTAAATTTGGACCTAGAAAAATATGTTGAT 879
Db 781 GTTCGAGAGCTATTAGAAGATGATGATTTAAACAATCATTTGGTCCAGAAAAATACATCGAT 840
Qy 880 ACTCAATATTTACCGCAGAACAAATTCCTTAACGATCCATATTTAATGAATACACTACCT 939
Db 841 ACACAAATATTTGACCCCAAAAGACTTCTTAATAAGCGAGTTTGTCTGGAATCAATTAACCA 900
Qy 940 GAAACCGCTACAAAATTAACATCTCTCGATTTACATCAAAAGGAAATATATCGTTGGATTGG 999
Db 901 GAAGTGAAACCAATAATAGTGTTCGCGCAAAAGGGGAGGAACAGTTTCTCTGGATTGG 960
Qy 1000 AGATTAGAACAATTTCAAAAAACCGATATCTACGCTATGTTGATCTCCGTTTCGTTAT 1059
Db 961 CGCTTTAGAACAAATTCGAAAAACAGAAAAATCTCCGCTTATCCGATTCGCTTTCCGCTTT 1020
Qy 1060 TTTGTTGCGGGTAATGTTGCAATTTTCTAAAGAAATGGCTAAATAAGTAGTTGGTTGTCGAT 1119
Db 1021 TTTGCGCGGGTAATGTTGCTTTCCGCTAAAAAATGGCTAAATAAGTAGTTGGTTGTCGAT 1080
Qy 1120 GAAGAAATTTAATTCATTTGGGGGCGGAAGATGTAGAAATTTGGTTACAGATTTATTTGCCAAA 1179
Db 1081 GAGGAATTTAATCACTCTGGGTGAGAGATGTGGAATTTGGATATCGCTTATTTCCGTTAC 1140
Qy 1180 GCCTGTTTTTCAGAGTAATTTGACGGGGAATGGCCATCCATCAAGAACCACTGGTAA 1239
Db 1141 GGTAGTTTTCTTTAAACATTTGATGGCATTTATGGCTACCATCAAGAGCCACAGGTAA 1200
Qy 1240 GAAATGAACAGACGCAAGCTGTGTAAGCTATTACGCTTAAATTTCTGTAAGAAAAAG 1299
Db 1201 GAAATGAACCGATCGGTGAGCGGGAATAATATTACCGCTCGATATTATGAGAGAAAAAG 1260
Qy 1300 GTACCTTACATCTATAGAAAGCTTTTACCAATAGAAGATTACATATTTCAAGAAATACCT 1359
Db 1261 GTCCCTTATCTATAGAAACCTTTTACCAATAGAAGATTTCGCATATCAATAGAGTACCT 1320
Qy 1360 TTAGTTTCTTATTTATATCCCGCTTATAACTGTGCAAAATATATTCAAAAGATGTGTAGAT 1419
Db 1321 TTAGTTTCAATTTATATCCAGCTTATAACTGTGCAAACTATATTCAACGTTGCGGTAGAT 1380
Qy 1420 AGTGTCTTATCAACACTGTTGTCGATCTCGAGGTTTGTATTGTAACCATGTTTCAACA 1479
Db 1381 AGTGCACGAATCAGACAGTGTGTTGATCTCGAGGTTTGTATTGTAACCATGTTTCAACA 1440
Qy 1480 GATAATACCTTTAGAAGTGTATCAATAAGCTTTTATGGTAAATAATCCTTAGGTAGGTCAGCATG 1539

[illegible]

FEATURE:
NAME/KEY: RBS
LOCATION: 340..345
US-08-746-682A-1

Query Match
Best Local Similarity 50.7%; Pred. No. 1e-07;
Matches 179; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 1343 ATATTCATAGAAATACCTTTAGTCTTATATATCCCGCTTATAACCTGTGCAAAATATA 1402
DB 8219 AAATGTATCTTAAAGTCTAATCTCTATTTGTTATTCAGTATATAATGTAGAGAAATATT 8278
QY 1403 TTCAAAGATGTAGATAGTCTTAAATCAACCTGTTGTCGATCTCGAGGTGTTGTAATT 1462
DB 8279 TAGAAAATGTTTGGCAATCTGTTCAAAATCAGACTTACAATAATTTTGAAGTGATTTAG 8338
QY 1463 GTAAGATGGTTCAACAGATAATACCTTTAGAGTGTATCAATAGCTTTATGGTAATATC 1522
DB 8339 TGAATGATGGCTCAACCGATTCATCACTTTCAATATGCGAAAAATTTGTTATCAGGATA 8398
QY 1523 CTAGGTAGCATCATGCTCAACCAATGCGGGAATAGCCCTCAGCATCAAAATGCGCGG 1582
DB 8399 AAAGATTTTCTGTTTCTTAAAGAAATGTTGGTGTGTCATCTGCACGAAATTTTGGNA 8458
QY 1583 TTTCTTTTCTTAAAGTATTACATTTGGGCAGTTAGATTTCAGATGATTATCTTTGAGCCCTG 1642
DB 8459 TTAAGAGCTAAGAGTGGTTTATCACATTTGTAGATAGTGTGATGACTACATAGTAAAG 8518
QY 1643 ATGCGATGGAATGCTTTTAAAGAAATTTTAAAGATATAAAGCGTAGCTTCT 1695
DB 8519 ATTATCTTCTCATTTGGTAGTGGGATAAAAGTGAGACCTCTATAGTTTGT 8571

RESULT 5
US-08-998-416-186/c
Sequence 186, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Rebischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBVA GOSSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:

NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 19390 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-86
Query Match 2.0%; Score 61; DB 4; Length 19390;
Best Local Similarity 48.7%; Pred. No. 0.00013;
Matches 166; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
Qy 1353 AATACCTTTAGTTCTATTTATATCCCGCTTATACTGTCGAATATATTTCAAGATG 1412
Db 4585 AACTGCAATTAATAGTGTGATTTGCGCAGTCTATAATGTGCGCAGTACCTAGAAAATC 4526
Qy 1413 TGTAGATAGTCTCTTAATCAAACTGTTGCGATCTCGAGGTTTGTATTTGTAACGATG 1472
Db 4525 GATAGCTTCCATTCAGAGCAGACCTATCAAAATCTGGAATTTATCTTGTGTATGATG 4466
Qy 1473 TTCACAGATTAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATATCCTAGGGTACG 1532
Db 4465 TGCACAGATGAAAGTGTGCTGCTTTGTGTATTCATCGCTGAACAAGATGACAGGGTGT 4406
Qy 1533 CATCATGCTAAACCAATGCGGAATAGCCTCAGCATCAATGCAAGCGGTTCTTTTTCG 1592
Db 4405 AGTGCTTCATAAAGACGAGGATGTCOCAGCAGCAATGATGGATGAAGCAGCG 4346
Qy 1593 TAAAGGTTATTACATTTGGCGAGTTAGATTCAGATCATTTATCTTGAGCCGTGATGAGTTGA 1652
Db 4345 TCACGGGATATCTGATTTTATTTTATGACTCAGATGATATATCCATCCAGAAATGATTCA 4286
Qy 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAACGCTACTT 1693
Db 4285 GAGCTTATGAGCAATTAGTTCAAGAAGATGCGGATGTTT 4245
RESULT 12
US-08-961-083-199
Sequence 199, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 199:
SEQUENCE CHARACTERISTICS:
LENGTH: 2023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-199
Query Match 2.0%; Score 60.6; DB 3; Length 2023;
Best Local Similarity 44.6%; Pred. No. 8.1e-05;
Matches 237; Conservative 0; Mismatches 294; Indels 0; Gaps 0;
Qy 1411 TGTGATAGTGTCTTAATCAAACTGTTGTCGATCTCGAGGTTTGTATTGTAACGAT 1470
Db 2 TGCCTAGATAGTATTATTACTCAACATATAAAATATTGAGATTGTTGTCGTTAATGAT 61
Qy 1471 GGTTCACAGATAATACCTTAGAAGTGATCAATAGCTTTATGTTAATAATCCTTAGGTA 1530
Db 62 GTTTCACGAGTCTTCAGTGAAATTTGTAAGAAATTTTCAGAAATGATGATCAGCAAT 121
Qy 1531 CGCATCATGCTTAAACCAATGCGGAATAGCCTCAGCATCAATGAGCCGTTCTTTT 1590
Db 122 CTCATATAGAACAAAGAAATGCTGCTCTTCTGCGCACGAAACACCGTCTGAATAAT 181
Qy 1591 GCTAAAGTTTATTACATTGGGCAGTTAGATTCAGATGATTATCTTGAGCCTGATGAGTT 1650
Db 182 ATGTCGGAATTTATGTGACCTTTGCGACTCGGATGATGATGATGACGAATGATGTA 241
Qy 1651 GAACCTGTCTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTTGTTTATACCACTAAT 1710
Db 242 GAACCTCTATATAAAATAAGTAGATGATCAGGCTGATATTGACGTTGGTAATATTAT 301
Qy 1711 AGAACGCTCAATCCGGATGGTAGCTTATTCGGTATGTTAGCTTACAAATGGCCACAATTTCA 1770
Db 302 TCTTTCAACGAAGTGAAGGAATGTTCTACTTTCTATATATTTGGGAGACTCTCTATTATGAG 361
Qy 1771 CGAGAAACACTCACACGCTATGATGCTCACCAATTTTAGAATGTTTACGATTAGAGCT 1830
Db 362 AAAGTATATGATAATGTTTCTATCTTTGAGAATCTGATGAACCTCAAGAAATGAGACT 421
Qy 1831 TGGCATTTAAACGGATGGATTTAACGAAATATTGAAACGCGCTGATGATGACATGTTTC 1890
Db 422 TTTGCTTTGATATCTGCTTGGGTAAACCTCTATAAGGCAAGATTTGTTTGGCAGTTGCGC 481
Qy 1891 CTTAACTCAGTGAAGTTGGAAATTTAAACATCTTAATAAATCTGCTAT 1941
Db 482 TTTGACATAGGTAAATTAGGAGAAGATGTTTACCTCAATCAAAAGGTATAT 532

```

; LOCATION: 3770...4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orig)" from
; ; OTHER INFORMATION: plasmid RP4"
; ;
US-09-056-075-1

Query Match 1.9%; Score 57.8; DB 2; Length 6243;
Best Local Similarity 45.0%; Pred. No. 0.00048;
Matches 300; Conservative 0; Mismatches 362; Indels 5; Gaps 2;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 23:56:15 ; Search time 168 Seconds
(without alignments)
7669.230 Million cell updates/sec

Title: US-09-842-484A-1
Perfect score: 2979
Sequence: 1 ttataaactgattaaagaag.....acattgcattttattaaaa 2379

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 381593 seqs, 216252194 residues
Total number of hits satisfying chosen parameters: 763186

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Published_Applications_NA:*	
Result No.	Score	Query Match	Length DB ID	Description	
1	2283.4	76.6	2937 9	US-09-879-959-9	Sequence 9, Appli
2	86	2.9	5550 10	US-09-070-927A-137	Sequence 137, App
3	80.6	2.7	6865 10	US-09-900-038A-3	Sequence 3, Appli
4	79.2	2.7	11474 10	US-09-816-028A-1	Sequence 1, Appli
5	74.2	2.5	17276 9	US-09-870-759-83	Sequence 83, Appli
6	73.2	2.5	6992 10	US-09-767-041-9	Sequence 9, Appli
7	73.2	2.5	6992 10	US-09-767-041-29	Sequence 29, Appli
8	68.4	2.3	891 10	US-09-816-028A-308	Sequence 30, Appl
9	67.2	2.3	912 10	US-09-816-028A-28	Sequence 28, Appl
10	66.8	2.2	1956 10	US-09-351-794A-1	Sequence 1, Appli
11	66.6	2.2	906 10	US-09-816-028A-26	Sequence 26, Appl
12	61.8	2.1	2406 10	US-09-070-927A-308	Sequence 308, App
13	60.6	2.0	2023 10	US-09-765-272-199	Sequence 199, App
14	58.8	2.0	516 10	US-09-960-352-5785	Sequence 5785, Ap
15	58.6	2.0	431 10	US-09-960-352-5558	Sequence 5558, Ap
16	57.6	1.9	811 10	US-09-765-272-197	Sequence 197, App
17	56	1.9	640681 10	US-09-790-988-1	Sequence 1, Appli
18	55	1.8	446 10	US-09-960-352-3400	Sequence 3400, Ap
19	55	1.8	3876 12	US-10-051-952-4	Sequence 4, Appli

c 20	53.8	1.8	519	10	US-09-878-574-4292	Sequence 4292, Ap
c 21	53.2	1.8	529	10	US-09-983-965-2109	Sequence 2109, Ap
c 22	52.6	1.8	479	10	US-09-960-352-12872	Sequence 12872, A
c 23	52.4	1.8	428	10	US-09-960-352-573	Sequence 573, App
c 24	52.2	1.8	1035	9	US-09-891-641-40	Sequence 40, Appl
c 25	51.2	1.7	414	10	US-09-960-352-6528	Sequence 6528, Ap
c 26	51	1.7	480	10	US-09-960-352-5301	Sequence 5301, Ap
c 27	50.6	1.7	377	10	US-09-960-352-5583	Sequence 5583, Ap
c 28	50.6	1.7	2012	10	US-09-070-927A-443	Sequence 443, App
c 29	50	1.7	408	10	US-09-960-352-1221	Sequence 1221, Ap
c 30	49.8	1.7	442	10	US-09-960-352-12911	Sequence 12911, A
c 31	49.8	1.7	1431	10	US-09-774-414-2	Sequence 2, Appli
c 32	49.6	1.7	640681	10	US-09-790-988-1	Sequence 1, Appli
c 33	49	1.6	2000	9	US-09-938-842A-4090	Sequence 4090, Ap
c 34	48.8	1.6	393	10	US-09-960-352-4582	Sequence 4582, Ap
c 35	48.6	1.6	4507	9	US-10-153-273-3	Sequence 3, Appli
c 36	48.6	1.6	5154	10	US-09-070-927A-201	Sequence 201, App
c 37	48.2	1.6	411	10	US-09-960-352-14521	Sequence 14521, A
c 38	48.2	1.6	910	10	US-09-765-272-201	Sequence 201, App
c 39	48	1.6	416	10	US-09-960-352-4584	Sequence 4584, Ap
c 40	48	1.6	40645	9	US-10-216-441-3	Sequence 3, Appli
c 41	48	1.6	40645	10	US-09-818-656A-3	Sequence 3, Appli
c 42	47.8	1.6	426	10	US-09-960-352-6529	Sequence 6529, Ap
c 43	47.6	1.6	601	10	US-09-818-656A-54	Sequence 54, Appl
c 44	47	1.6	376	10	US-09-960-352-5087	Sequence 5087, Ap
c 45	47	1.6	1287	9	US-09-286-488-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-879-959-9
; Sequence 9, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kghama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISITMILIS HYALURONAN SYNTHASE GENE AND EXPRES
; TITLE OF INVENTION: IN BACILLUS SUBTILIS
; FILE REFERENCE: 3554.049
; CURRENT FILING DATE: 2001-09-12
; PRIOR FILING DATE: 1999-12-21
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: pasteurella multocida
US-09-879-959-9

Query Match	76.6%	Score 2283.4	DB 9	Length 2937
Best Local Similarity	86.5%	Pred No. 0		
Matches 2540	Conservative 0	Mismatches 376	Indels 21	Gaps 1
Qy 43	ATTTTAAAGGAAGAAATGAATATATCACAAGCAATAAAGCATATACAGCAAT	102		
Db 1	ATTTTAAAGGACAGAAATGAATATATCACAAGCAATAAAGCATATACAGCAAT	60		
Qy 103	GACTATGAATAGCACTCAAAATTTTGAAGAGTCTGCTGAACCTACGGCGGCAAAATC	162		
Db 61	GACTATCAATTAGCACTCAAAATTTTGAAGAGTCTGCTGAACCTACGGCGGCAAAAT	120		
Qy 163	GTGTAATTCACCAATTTATCAAAATGTAAGAAAAAATCTCCAGCAATTC-----	209		
Db 121	GTGTAATTCACCAATTTATCAAAATGTAAGAAAAAATCTCCAGCAATTCCTTCTGTAAATTC	180		
Qy 210	-----TTATGTAAGTGAAGATAAAAAAAGAGTGTGTTGGGATAGCTCATATAGATATC	261		

Db	181	GCATCTTTCTGTGTAATAAAGAGAAAAGCTCAATGTTTCGGATAGTCCGTTAGATATT	240
Qy	262	GCAACACAGCTCTTACTTTTCCAAAGTAAAAAAATTAACATCTATCCGAATCAGAAAAAAC	321
Db	241	GCAACAACTGTACTTTTCCAAAGTAAAAAAATTAAGTACTTCTGACTCGGAAAAAAC	300
Qy	322	AGTTTAAAAAATAATGGAATCTATCACTCGGAAAAAATCGGAGAACGCGAGAAATCAGA	381
Db	301	ACGTTAAAAAATAATGGAATTCGCTACTGAGAGAAATCTGAATAATCGGAGGTAAGA	360
Qy	382	AAGTGGAACTAGTACCCAAAGATTTTCCCTAAAGATCTTGTTCTTGCTTCCATTTGCCAGAT	441
Db	361	CGGTGCGCCCTGTACCAAAAAGATTTTCCCAAGATCTGGTTTATAGCCCTTTACCTGAT	420
Qy	442	CATGTTAATGATTTTACATGTTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCTGTGA	501
Db	421	CATGTTAATGATTTTACATGTTACAAAAAGGAGAAAGAACTTGGCATTAACCTGAA	480
Qy	502	AATAGAAATACGGTCTTCTTCTATTATTATTCCTACATTTAATCGTAGCGTATTTTAGAT	561
Db	481	CATCAACATGTTGGTCTTCTTATTATCGTTTACAACATTCATTCGACACGCAATTTTATCG	540
Qy	562	ATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTACCCATTTGAAGTCGTTGTGCA	621
Db	541	ATTACATTAGCCTGTTTAGTAACCAAAAAACACATTACCCGTTTGAAGTTATCGTGACA	600
Qy	622	GATGATGTTAGTAAGAAAACTTACTTACCATTGTGCAAAAATACGACAAAAAAGTTGAC	681
Db	601	GATGATGTTAGTCAGGAAGATCTATACCGGATCATTCGCCAATATGAAATAAATTTGGAT	660
Qy	682	ATAAGTATGTAAGACAAAAAGATTATGGATATCAATTGTGTGACGTACGAACTTTAGT	741
Db	661	ATTCTGCTAGTACAGAAAAAGATACGGTTTTCAAGCCAGTTCGCGCTCGGAATATGGGA	720
Qy	742	TTACGTCACAGAAAGTATGATTTTGTCTCGATTCTAGACTGCGATATGCGACCAACAA	801
Db	721	TTACGCTTAGCAAAATATGACTTTATTGGCTTACTCGACTGTATGCGCAAAATCCA	780
Qy	802	TTATGGGTTCACTCTTATCTTACAGAACTATTAGAAGAACTATGATTTGTTTAAATTGA	861
Db	781	TTATGGGTTCACTCTTATCTTGCAGAGCTATTAGAAGATGATGATTAAACAATCATTTGT	840
Qy	862	CCTAGAAATATGTGATACTCATATATTACCGCAGAACAAATTCCTTAAAGATCCATAT	921
Db	841	CCAAGAAATATCATGATACACAACTATTGACCCAAAAGACTCTTAAATAACGCGAGT	900
Qy	922	TTAATAGAACTACTTACCTGAAACCGCTACAAAATAACAATCCTTCGATTACATCAAAAGGA	981
Db	901	TTGCTTGTAATCATTACCAGAGTGAACCAATAATAGTTGTTCCGCAAAAGGGGAAGGA	960
Qy	982	AATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAACCGATAATCTACGCTCATGT	1041
Db	961	ACAGTTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAACAGAAATCTCCGCTTATCC	1020
Qy	1042	GATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTTTGCATTTTCTAAAGAAATGGCTAAAT	1101
Db	1021	GATTCGCCCTTCGGTTTTTTTGGCGCGGTAAATGTTTGCCTTAAAAAATGGCTAAAT	1080
Qy	1102	AAAGTAGGTTGGTTTCGATGAAGAATTTAATCATTTGGGGGGCGGAAGATCTAGAAATTTGGT	1161
Db	1081	AAATCCGGTTCTTTCATGAGGAATTTAATCACTGGGTGGAGAAGATGTGGAATTTGGA	1140
Qy	1162	TACAGATTAATTTGCCAAAGCTGTTTTTTCAGAGTAATTGACGGCGGAATGGCCATCCAT	1221
Db	1141	TATCGCTTATTCGGTTACGGTAGTTTCTTTTAAACATATTGATGGCATTTATGGGCTACCAT	1200
Qy	1222	CAGAACCACCTCGTAAAGAAAAATGAACAGAACGCGAAGCTGGTAAAGATTTACGCTT	1281
Db	1201	CAGAGCCACCGGTTAAAGAAAAATGAACCGATTCGTGAACGGGGAAAAAATATTACGCTC	1260
Qy	1282	AAAATTTGTAAGAAAGAGTACCTTACATCTATAGAAAAGCTTTTACCAAATAGAAAGATTCA	1341


```
; SEQ ID NO 30
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(891)
; OTHER INFORMATION: beta-1,3 galactosyl transferase from C. jejuni O:10
US-09-816-028A-30

Query Match          2.3%; Score 68.4; DB 10; Length 891;
Best Local Similarity 52.8%; Pred. No. 0.00011;
Matches 172; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

QY 1362 AGTTCTATTATATCCCGCTTATAACTGTGCAAAATTTATATCAAGATGTTGTAGATAG 1421
   ||||| || || ||||| || || ||||| || || ||||| || || ||||| || || |||||
Db 9 AATTTCATCATACTACCAACTTATAATGTGGAAAAATATATGCTAGAGCATTAGAAG 68

QY 1422 TGCCTCTTAATCAAACTGTTGTCGATCTCGAGGTTTGTATTTGTAAGCATGTTCAACAGA 1481
   ||||| || || ||||| || || ||||| || || ||||| || || ||||| || || |||||
Db 69 TTGCATTAACCAAACTTTTAAAGATATAGAAAATCATTTGTAGTAGATGTTGGTAATGA 128

QY 1482 TAATACCTTAGAAGTGATCAATAAGCTTTATGCTAATAATCCCTAGGTCAGCATCATGTC 1541
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 129 TAAAGTATAGATATAGCTTAAGAGATGCTAGTAAGATGATAGATAAAATCATACA 188

QY 1542 TA---AACCAAAATGGGGAATAGCCTCAGCATCAAAATGCAGCCGTTTCTTTTGTCTAAAGG 1598
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 189 TAATGAAGAAATTTAAAGCTTTTAAAGCAAGATATGAAGGTGCTAAAGTAGCAACTTC 248

QY 1599 TTATTACATGGGCAGTTAGATTGATGATGATTTATCTTGAGCCCTGATGCAGTTGAACTGTG 1658
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 ACCTTATCATGCTTTTGTAGATTCTGATGATTTTATGAACCTTTAATGCTTGCAGAAGATG 308

QY 1659 TTTTAAAGAATTT 1670
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 TATTAAATTTT 320
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-351-794A-1
; Sequence 1, Application US/09351794A
; Patent No. US20020042382A1
; GENERAL INFORMATION:
; APPLICANT: DUFFY, PATRICK E.
; APPLICANT: OCKENHOUSE, CHRISTIAN F.
; TITLE OF INVENTION: SEQUESTRAIN
; FILE REFERENCE: 38644-175519
; CURRENT APPLICATION NUMBER: US/09/351.794A
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 08/559,896
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
US-09-351-794A-1

Query Match          2.2%; Score 66.8; DB 10; Length 1956;
Best Local Similarity 45.8%; Pred. No. 0.00028;
Matches 273; Conservative 0; Mismatches 317; Indels 6; Gaps 1;

QY 1987 ATTCAAAAGAAAAACCACTTTTGTGTAGTCAATCAGTCATTAATAGACAGGCAATCAAT 2046
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1228 ATACAGCAATAATGATAGTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1287

QY 2047 TATTATAATTATGACAAATTTGATGATTTAGATGAAAGTAGAAGTATATCTTCAATAAA 2106
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1288 TTAATAATAAGCACACAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1347

QY 2107 ACCGCTGAATATCAAGAAGAAATGGATATGTTTAAAGATCTTTAAACTCATTCACAAATAAA 2166
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1348 CGAAATTTAGCAAGGATAAAGGAGGCAAGATAGAGATATATATAGACTATTTTAAACAAA 1407

QY 2167 GATGCCAAATCGCAGTCAGTATTTTCTTATCCCATACATTAACGGCTTAGTGAATAAA 2226
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1408 GAAATTTAAAGAACAAAAATGTTAATGTTCCCAATATAGTAAGTATTTTAAATTCAAAA 1467

QY 2227 CTAAACAATATTGTAATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2286
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1458 GTAGGAAGAGATAACACCACTTCAACATAAGAAGGAAATCAAGTAGATGTTGTGAGG 1527

QY 2287 AATCATCTTACACAGACATCAAAAAAGAAATATTGGCGTTCTTATCATCAAGCACAAGTG 2346
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
; SEQ ID NO 30
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(891)
; OTHER INFORMATION: beta-1,3 galactosyl transferase from C. jejuni O:10
US-09-816-028A-30

Query Match          2.3%; Score 68.4; DB 10; Length 891;
Best Local Similarity 52.8%; Pred. No. 0.00011;
Matches 172; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

QY 1362 AGTTCTATTATATCCCGCTTATAACTGTGCAAAATTTATATCAAGATGTTGTAGATAG 1421
   ||||| || || ||||| || || ||||| || || ||||| || || ||||| || || |||||
Db 9 AATTTCATCATACTACCAACTTATAATGTGGAAAAATATATGCTAGAGCATTAGAAG 68

QY 1422 TGCCTCTTAATCAAACTGTTGTCGATCTCGAGGTTTGTATTTGTAAGCATGTTCAACAGA 1481
   ||||| || || ||||| || || ||||| || || ||||| || || ||||| || || |||||
Db 69 TTGCATTAACCAAACTTTTAAAGATATAGAAAATCATTTGTAGTAGATGTTGGTAATGA 128

QY 1482 TAATACCTTAGAAGTGATCAATAAGCTTTATGCTAATAATCCCTAGGTCAGCATCATGTC 1541
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 129 TAAAGTATAGATATAGCTTAAGAGATGCTAGTAAGATGATAGATAAAATCATACA 188

QY 1542 TA---AACCAAAATGGGGAATAGCCTCAGCATCAAAATGCAGCCGTTTCTTTTGTCTAAAGG 1598
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 189 TAATGAAGAAATTTAAAGCTTTTAAAGCAAGATATGAAGGTGCTAAAGTAGCAAACTC 248

QY 1599 TTATTACATGGGCAGTTAGATTGATGATGATTTATCTTGAGCCCTGATGCAGTTGAACTGTG 1658
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 TCCTTATATATGTTTGTAGCTGATGATTTTATGAACCTTTAATGCTTGTGAAGAATG 308

QY 1659 TTTTAAAGAATTTTAAAGATGAAA 1684
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 TATGAAAAATTTTAAAGAAATGAAA 334
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-09-816-028A-28
; Sequence 28, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816.028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 912
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(912)
; OTHER INFORMATION: Campylobacter glycosyltransferase B (CgtB) beta-1,3
; OTHER INFORMATION: galactosyltransferase from C. jejuni serotype O:2
; OTHER INFORMATION: (strain NCIC 11166)
US-09-816-028A-28

Query Match          2.3%; Score 67.2; DB 10; Length 912;
Best Local Similarity 53.2%; Pred. No. 0.0002;
Matches 166; Conservative 0; Mismatches 143; Indels 3; Gaps 1;
```

Db 1528 AAAAATATTCAGATATTTCAAGAGGATATATATAAAAAATAAAGCCAAAAGGATAAACACT 1587
Qy 2347 AATATTTTACTAAATATGACATCTCATATTTACACGAGTATACAGTAATAAAAACTGAG 2406
Db 1588 GAAATGTTAGATATATATAGGAATA------ACAATATTTGATATAAAAATTTGAT 1641
Qy 2407 GCACATTTAAGTAAATTAATTAATTAAGTCAGTTAAATCTTAAATCTTGAATACATCAT 2466
Db 1642 GATATAAAAAATGTTGGAGATATAAAAAAGTTTCAGATATAAAAAAGTTTGGATGATATA 1701
Qy 2467 TTTGATATCATGACACGCTATTCGTTTAAATATCACAGCTATGCTTATATGAAAAATAT 2526
Db 1702 AACAACTTTGATGCTATATAAAAAATGTTTGGATGATATAAAAAATGTTGATGTTAAAAAT 1761
Qy 2527 GATCTGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAATGC 2582
Db 1762 GTTGATGCTATAAACAATGTGGGAGATATAAACAATGCTGGAGATACAAATAATGC 1817

RESULT 11
US-09-816-028A-26
; Sequence 26, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
; OTHER INFORMATION: beta-1,3-galactosyltransferase from C. jejuni strain
; OTHER INFORMATION: OH4384 (ORF 6a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus
US-09-816-028A-26

Query Match 2.2% Score 66.6; DB 10; Length 906;
Best Local Similarity 51.3%; Pred. No. 0.00026;
Matches 181; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

Qy 1362 AGTTTCTATTTATATCCCGCTTATAACTGTGCAAAATATATATCAAGATGTGTAGATAG 1421
Db 9 AATTTCAATCATCTTACCAACTTATATGTGGAACAATATATACGAAGGCAATAGAAG 68
Qy 1422 TGTCTTAAATCAAACTGTTGTCGATCTCGAGGTTTGTATTTGTAAACGATGTTCACAGA 1481
Db 69 CTGTATCAATCAGACTTTTAAAGATATAGAAATAATTTAGTATGATGTTGGAAATGA 128
Qy 1482 TAATACCTTAGAAGTATCAATAGCTTTATGTTAATTAATCTAGGGTACGCATCATGTC 1541
Db 129 TAATAGTATAATATAGCCAAGGAATACTCTAAAAGACAAAAAATAAATAATCA 188
Qy 1542 TAAACCAAAATGCC---GGAATAGCTCAGCAATCAAAATGACGCCGTTTCTTTTGTAAAGG 1598
Db 189 CAATGAAAAAAACTTAGGTCTTTTAAAGGACGAAGATATGAAGTGTGAAGTAGCAACTC 248
Qy 1599 TTATTACATTTGGGAGTTAGATTCAGATGATTATCTTGAGCGCTGATGCAAGTTGAACTGTG 1658
Db 249 TCCTTATATAATGTTTTAGATCTCTGATGATTATTTTGGAACTAAATGCTTGTGAAGACTG 308

Qy 1659 TTTAAAGCAATTTTAAAGATAAAACGCTAGCTTGTGTATTATACCACATAA 1711
Db 309 TATAAAATTTTAGATGAACAGGATGAAGTTGATTAGTGTGTTTTCATGCTA 361

RESULT 12

US-09-070-927A-308/c
; Sequence 308, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 382
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2406 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 308:
US-09-070-927A-308

Query Match 2.1% Score 61.8; DB 10; Length 2406;
Best Local Similarity 50.0%; Pred. No. 0.0028;
Matches 147; Conservative 1; Mismatches 146; Indels 0; Gaps 0;

Qy 1371 TTATATCCCGCTTATAACTGTGCAAAATATATATCAAGATGTGTAGATAGTCTCTTAA 1430
Db 2311 TTGTNCCNGCTCTATTAACTGGAACCCCTATTTTGGNGGAGGCATTAATCAGTTTGA 2252
Qy 1431 TCAAACTGTTGTCGATCTCGAGGTTTGTATTTGTTAACTGTTTCAACAGATTAATACCT 1490
Db 2251 CCAAACTATGAAGAATTTGAATTTCTTAATTAATGATGCTCTACTGATAATAGTCA 2192
Qy 1491 AGAAGTGTCAATAGCTTTATGGTAAATATCTAGGGTACGCATCATGCTTAAACCAA 1550
Db 2191 AAAAATTTGAAGAAATTCACAAGATGATCCAGCTTTTCGCCGTTTCAATGTTAAAA 2132
Qy 1551 TGGCGGAATAGCTTCAGCATCAATGCACCGCTTTCTTTTGTAAAGGTTATTACATTGG 1610
Db 1551 TGGCGGAATAGCTTCAGCATCAATGCACCGCTTTCTTTTGTAAAGGTTATTACATTGG 1610

Db 154 AAAAAAAAAAATAAAAAATAAAAAAATAAAAATTAAACAAAACA 95
Qy 2445 TCTAAATTGCGAATACATCATTTTTTGATATCATGA 2480
Db 94 ATATNAAAGTAACCCCTCNATACCTTATTGGCATT 59

RESULT 15

```

RESULT 15
US-09-960-352-5558/c
; Sequence 5558, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wyatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5558
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 24-LIB3057-024-Q1-K1-F7
US-09-960-352-5558

```

Query Match	2.0%;	Score 58.6;	DB 10;	Length 431;	
Best Local Similarity	47.6%;	Pred. No. 0.0076;			
Matches 172;	Conservative 0;	Mismatches 189;	Indels 0;	Gaps 0;	
Qy	2104	AAAAACCGCTGAATATCAAGAGGAATGATGATGTTAAAAAGATCTTAAACTCATTTCAAAT	2163		
Db	431	AAAAAAAAATAAAAAAAAAATAAAAAAAAAATAAAAAAAAAATAAAAAAAAAATAAAAAAAAAAT	372		
Qy	2164	AAAGATGCCAAAATCGCAGTCAGTATTTCTTATCCCAATACATTTAAACGGCTTAGTGCAA	2223		
Db	371	TAATAAAAAATAAATAAATAAAAAATATAAAAAAATAAAAAAATAAAAAAATAAAAAATAA	312		
Qy	2224	AAACTAAACAATATTATTGAAATATAATAAAAAATATATTCGTATTATCTACATGTTGAT	2283		
Db	311	ATAATAAAAAATAAAAAACAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA	252		
Qy	2284	AAGAACTCATCTTACACAGACATCAAAAAAGAAATATTGGCTTCTATCATTAAGCACCAA	2343		
Db	251	AAAAAAATTAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATAATAA	192		
Qy	2344	GTGAATATTTTACTAATAATAGACATCTCATATTATACAGGAGTAATAGACTAATAAAAACT	2403		
Db	191	AAAAAAATAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAT	132		
Qy	2404	GAGGCACATTTAAGCTAAATATTATAATAATTAAGTCAGTTAAATCTAAATTTGAATACATC	2463		
Db	131	ATAAAAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATAAATTAATTTATAGATA	72		
Qy	2464	A	2464		
Db	711	A	71		

Search completed: January 4, 2003, 02:22:36
Job time : 277 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:20:09 ; Search time 5768 Seconds
(without alignments)
12985.345 Million cell updates/sec

Title: US-09-842-484a-1
Perfect score: 2979
Sequence: 1 ttataaactgattaaagaag.....acatttgcattttattataaaa 2979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 24791104 seqs, 12571243825 residues 49582208
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/pctus_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2979	100.0	2979	1	PCT-US01-1335-1 Sequence 1, Appl
2	2979	100.0	2979	1	PCT-US99-2650-4 Sequence 4, Appl1
3	2963	99.5	2979	1	PCT-US01-1335-3 Sequence 3, Appl1
4	2285	76.7	2937	38	US-10-011-766B-8 Sequence 8, Appl1
5	2285	76.7	2937	38	US-10-011-771B-8 Sequence 8, Appl1
6	2283.4	76.6	2937	1	PCT-US99-07289-2 Sequence 2, Appl1
7	2283.4	76.6	2937	16	US-09-283-402-2 Sequence 2, Appl1
8	2283.4	76.6	2937	18	US-09-469-200-2 Sequence 2, Appl1
9	2283.4	76.6	2937	33	US-09-879-959-9 Sequence 9, Appl1
10	2283.4	76.6	2937	41	US-10-172-527-9 Sequence 9, Appl1
11	1503.2	50.5	2112	1	PCT-US99-26501-2 Sequence 2, Appl1
12	89.2	3.0	984	15	US-09-107-532-2527 Sequence 2527, Ap
13	89.2	3.0	984	15	US-09-107-532A-2527 Sequence 2527, Ap
14	86	2.9	993	15	US-09-134-000-2987 Sequence 2987, Ap
15	86	2.9	5550	14	US-09-070-927-137 Sequence 137, App
16	86	2.9	5550	14	US-09-070-927A-137 Sequence 137, App
17	82.2	2.8	908	26	US-09-666-355A-6574 Sequence 6574, Ap
18	82	2.8	1368	38	US-10-015-127-2421 Sequence 2421, Ap
19	80.6	2.7	6865	34	US-09-900-038A-3 Sequence 3, Appl1
20	79.4	2.7	1095	38	US-10-015-127-1938 Sequence 1938, Ap
21	79.2	2.7	11474	18	US-09-495-406-1 Sequence 1, Appl1

c 22 79.2 2.7 11474 31 US-09-816-028-1 Sequence 1, Appli
c 23 79.2 2.7 11474 31 US-09-816-028A-1 Sequence 1, Appli
c 24 78.8 2.6 810 26 US-09-666-355A-6400 Sequence 6400, Ap
c 25 78.8 2.6 1002 15 US-09-107-532-1469 Sequence 1469, Ap
c 26 78.8 2.6 1002 15 US-09-107-532A-1469 Sequence 1469, Ap
c 27 77.6 2.6 577 33 US-09-874-708A-11415 Sequence 11415, A
c 28 77.6 2.6 577 65 US-60-211-750-11272 Sequence 11272, A
c 29 76.6 2.6 13825 24 US-09-634-238-29 Sequence 29, Appl
c 30 76.2 2.6 1141 31 US-09-806-708A-22 Sequence 22, Appl
c 31 76 2.6 7458 1 PCT-US99-26796-239 Sequence 239, App
c 32 75.8 2.5 777 17 US-09-397-761A-2730 Sequence 2730, Ap
c 33 75.8 2.5 777 32 US-09-838-601-2730 Sequence 2730, Ap
c 34 75.8 2.5 980 26 US-09-666-355A-6399 Sequence 6399, Ap
c 35 75.8 2.5 1776 38 US-10-015-127-2386 Sequence 2386, Ap
c 36 75.2 2.5 824 64 US-60-207-458-18907 Sequence 18907, A
c 37 75.2 2.5 1250 29 US-09-739-449-1859 Sequence 1859, Ap
c 38 75.2 2.5 1250 31 US-09-803-110-1859 Sequence 1859, Ap
c 39 75.2 2.5 18585 50 US-60-068-175-605 Sequence 605, App
c 40 75 2.5 581 38 US-10-020-338-12117 Sequence 12117, A
c 41 75 2.5 961 56 US-60-126-265-3039 Sequence 3039, Ap
c 42 74.8 2.5 948 18 US-09-404-520-13069 Sequence 13069, A
c 43 74.8 2.5 4985 39 US-10-056-405-10 Sequence 10, Appl
c 44 74.8 2.5 4985 39 US-10-094-240-10 Sequence 10, Appl
c 45 74.2 2.5 17276 33 US-09-870-759-83 Sequence 83, Appl

ALIGNMENTS

RESULT 1
PCT-US01-13395-1
; Sequence 1, Application PC/TUS0113395
; GENERAL INFORMATION:
; APPLICANT: DE ANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND
; FILE OF INVENTION: USING SAME
; FILE REFERENCE: 618755-9/JP/199,538
; CURRENT APPLICATION NUMBER: PCT/US01/13395
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2979
; TYPE: DNA
; ORGANISM: Pasteurella multocida
PCT-US01-13395-1

Query Match 100.0%; Score 2979; DB 1; Length 2979;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAACTGATTAAAGAGGTAACGATTCACAGCAAGGTTAAATTTTAAAGAAAGAAA 60
Db 1 TTATAACTGATTAAAGAGGTAACGATTCACAGCAAGGTTAAATTTTAAAGAAAGAAA 60
QY 61 ATGAATACATATACACAGCAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
Db 61 ATGAATACATATACACAGCAATAAAGCATATAACAGCAATGACTATGAATTAGCACTC 120
QY 121 AAATTTATTTGAGAAGTCTGCTGAAACCTACGGCGGCAAAAATCGTTGAATTCCAAATATC 180
Db 121 AAATTTATTTGAGAAGTCTGCTGAAACCTACGGCGGCAAAAATCGTTGAATTCCAAATATC 180
QY 181 AAATGTAAGAAACACTCTCGACCAATCTTATGTAAGTGAAGATATAAAACACAGTGT 240
Db 181 AAATGTAAGAAACACTCTCGACCAATCTTATGTAAGTGAAGATATAAAACACAGTGT 240
QY 241 TGGGATAGCTCATTAGATATCGCAACACAGCTCTTCTTCCACAGCTAAACAAATTAAC 300
Db 241 TGGGATAGCTCATTAGATATCGCAACACAGCTCTTCTTCCACAGCTAAACAAATTAAC 300

Db 241 TCGGATAGCTCATTAGATATCGCAACACAGCTCTTCTTCCACAGCTAAACAAATTAAC 300
QY 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGAAATCTATCACTGGGAAAAA 360
Db 301 CTATCCGAATCAGAAAAACAGTTTAAAAATAAATGAAATCTATCACTGGGAAAAA 360
QY 361 TCGGAGAACGCAAGAAATCAGAAAGGTGGAACCTAGTACCACAAAGATTTTCTTAAGATCTT 420
Db 361 TCGGAGAACGCAAGAAATCAGAAAGGTGGAACCTAGTACCACAAAGATTTTCTTAAGATCTT 420
QY 421 GTTCTTGCTCCATTGCGCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAA 480
Db 421 GTTCTTGCTCCATTGCGCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAA 480
QY 481 AGCTTAGGTATAAAGCCTGTAAATAAGATATACGGTCTTTCTATTATATTCCTACATTT 540
Db 481 AGCTTAGGTATAAAGCCTGTAAATAAGATATACGGTCTTTCTATTATATTCCTACATTT 540
QY 541 AATCGTAGCGGTATTTTAGATATACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTAC 600
Db 541 AATCGTAGCGGTATTTTAGATATACGTTAGCCTGTTTGGTCAATCAGAAAAACAACTAC 600
QY 601 CCATTTGAAGTCGTTGTCGAGATGATGGTAGAAGGAAAACTTACTTACCATTGTCAA 660
Db 601 CCATTTGAAGTCGTTGTCGAGATGATGGTAGAAGGAAAACTTACTTACCATTGTCAA 660
QY 661 AAATACGAAACAAAACTTGACATAAAGTATGTAAAGAAAAAAGATATGAGATATCAATTG 720
Db 661 AAATACGAAACAAAACTTGACATAAAGTATGTAAAGAAAAAAGATATGAGATATCAATTG 720
QY 721 TGTGCACTCAGAACTTAGTTTACGTACAGCAAGCTATGATTTTGTCTCGATTCTAGAC 780
Db 721 TGTGCACTCAGAACTTAGTTTACGTACAGCAAGCTATGATTTTGTCTCGATTCTAGAC 780
QY 781 TGGCATATGCGACCAACAACTTATGGTTCATCTTCTTACAGAACTATTTAGAAAG 840
Db 781 TGGCATATGCGACCAACAACTTATGGTTCATCTTCTTACAGAACTATTTAGAAAG 840
QY 841 AATGATATTTTAAATGGACCTAGAAAAATATGTTGGATCTCATATATATACCGCAGAA 900
Db 841 AATGATATTTTAAATGGACCTAGAAAAATATGTTGGATCTCATATAATATACCGCAGAA 900
QY 901 CAATTCCTTAACGATCCATATTTAATAGATCACTACTGAAACCGCTACAAATAACAAT 960
Db 901 CAATTCCTTAACGATCCATATTTAATAGATCACTACTGAAACCGCTACAAATAACAAT 960
QY 961 CCTTCGATTACATAAAGGAAATATATCGTTGGATTGGAGATTAGAACTTCAAAAAA 1020
Db 961 CCTTCGATTACATAAAGGAAATATATCGTTGGATTGGAGATTAGAACTTCAAAAAA 1020
QY 1021 ACCGATATCTACGCTATGATGATTCCTCGTTTCGTTATTTTGTGGGGTAATGTTCCA 1080
Db 1021 ACCGATATCTACGCTATGATGATTCCTCGTTTCGTTATTTTGTGGGGTAATGTTCCA 1080
QY 1081 TTTTCTAAAGAAATGCGTAAATAAAGTAGTTGGTTCGATGAAGAAATTAATCAATTTGGGG 1140
Db 1081 TTTTCTAAAGAAATGCGTAAATAAAGTAGTTGGTTCGATGAAGAAATTAATCAATTTGGGG 1140
QY 1141 GCGGAAGATGTAGAAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
Db 1141 GCGGAAGATGTAGAAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
QY 1201 GACGGCGGAATGGCCATCCATCAGAACCCACCTGGTAAAGAAATGAACAGACGCCAA 1260
Db 1201 GACGGCGGAATGGCCATCCATCAGAACCCACCTGGTAAAGAAATGAACAGACGCCAA 1260
QY 1261 GCTGTTAAAGTATTACGCTTAAATTTGAAAGAAAAAGGTACCTTACATCTATAGAAG 1320
Db 1261 GCTGTTAAAGTATTACGCTTAAATTTGAAAGAAAAAGGTACCTTACATCTATAGAAG 1320
QY 1321 CTTTACCAATAGAAAGTACATTTACATAGATACCTTTAGTTTCTTATTATATATCCCC 1380
Db 1321 CTTTACCAATAGAAAGTACATTTACATAGATACCTTTAGTTTCTTATTATATATCCCC 1380

Qy 1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTAGATAGTCTCTTAATCAAACTGTT 1440
Db 1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTAGATAGTCTCTTAATCAAACTGTT 1440
Qy 1441 GTCGATCTCGAGGTTTGTATTTGTAACAGATGGTTTCAACAGATAATACCTTTAGAAGTATC 1500
Db 1441 GTCGATCTCGAGGTTTGTATTTGTAACAGATGGTTTCAACAGATAATACCTTTAGAAGTATC 1500
Qy 1501 AATAAGCTTTATGCTAATAATCTTAGGTAGGCATCATGTCTAAACCAATGGCGGAATA 1560
Db 1501 AATAAGCTTTATGCTAATAATCTTAGGTAGGCATCATGTCTAAACCAATGGCGGAATA 1560
Qy 1561 GCCTCAGCATCAAAATGACGCGCTTCTTTTGTCTAAAGGTTATTACATTTGGGCGAGTTAGAT 1620
Db 1561 GCCTCAGCATCAAAATGACGCGCTTCTTTTGTCTAAAGGTTATTACATTTGGGCGAGTTAGAT 1620
Qy 1621 TCAGATGATTATCTTTGAGCGCTGATGCACTGAGTTGAAGTGTGTTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATTATCTTTGAGCGCTGATGCACTGAGTTGAAGTGTGTTTAAAGAAATTTTAAAGAT 1680
Qy 1681 AAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCAATCCGGATGGTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTGTGTTTATACCACTAATAGAAACGTCAATCCGGATGGTAGCTTAATC 1740
Qy 1741 GCTAATGGTTACAAATGGCCAGAAATTTTACGAGAAAAAATCAACACGGCTATGATTGCT 1800
Db 1741 GCTAATGGTTACAAATGGCCAGAAATTTTACGAGAAAAAATCAACACGGCTATGATTGCT 1800
Qy 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTTAAACGGATGGATTTTAAAGAAAT 1860
Db 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTTAAACGGATGGATTTTAAAGAAAT 1860
Qy 1861 ATTCAAACCGCGGATGATGACATGTTCTTAACTCAGTGAAGTTGGAAATTTTAA 1920
Db 1861 ATTCAAACCGCGGATGATGACATGTTCTTAACTCAGTGAAGTTGGAAATTTTAA 1920
Qy 1921 CATCTTAATAAAATCTGCTATAACCGGGATTTACATGGTGATAACACATCCATTAAGAAA 1980
Db 1921 CATCTTAATAAAATCTGCTATAACCGGGATTTACATGGTGATAACACATCCATTAAGAAA 1980
Qy 1981 CTCGGCATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAAGTAAAGCAAGGC 2040
Db 1981 CTCGGCATTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAATTAAGTAAAGCAAGGC 2040
Qy 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGATGAAAGTAAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGATGAAAGTAAAGTATATCTTC 2100
Qy 2101 AATAAAACCGCTGAATATCAAGAAGAAATGGATATGTTTAAAGATCTTAAACTCATTCAA 2160
Db 2101 AATAAAACCGCTGAATATCAAGAAGAAATGGATATGTTTAAAGATCTTAAACTCATTCAA 2160
Qy 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Db 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Qy 2221 AAAAACTAAACATATATTGAATATAATAAAATATTTCGTTATTATTCATCATGTT 2280
Db 2221 AAAAACTAAACATATATTGAATATAATAAAATATTTCGTTATTATTCATCATGTT 2280
Qy 2281 GATAAGATCATCTTACACCAAGACATCAAAAAGAAATATTGGCTTCTTATCATTAAGCAC 2340
Db 2281 GATAAGATCATCTTACACCAAGACATCAAAAAGAAATATTGGCTTCTTATCATTAAGCAC 2340
Qy 2341 CAAGTGAATATTTTACTAAATTAATGACATCTCATATTACAGGTAATAGACTAATAAAA 2400
Db 2341 CAAGTGAATATTTTACTAAATTAATGACATCTCATATTACAGGTAATAGACTAATAAAA 2400
Qy 2401 ACTGAGGCACATTTAAGTAATATTAAATTAAGTCAAGTTAAATCTAAATTTGTAATAC 2460
Db 2401 ACTGAGGCACATTTAAGTAATATTAAATTAAGTCAAGTTAAATCTAAATTTGTAATAC 2460

Qy 2461 ATCATTTTTTGATATCATGACAGCCCTATTTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Db 2461 ATCATTTTTTGATATCATGACAGCCCTATTTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Qy 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTTAAACACATGATTGGATCGAGAAAAATCAAT 2580
Db 2521 AAATATGATGTCGGCATGAATTTCTCAGCATTTAAACACATGATTGGATCGAGAAAAATCAAT 2580
Qy 2581 CGCATCCACCATTTTAAAAAGCTGATTAAAACTTATTTTAAATGACAATGACTTAAAGAAGT 2640
Db 2581 CGCATCCACCATTTTAAAAAGCTGATTAAAACTTATTTTAAATGACAATGACTTAAAGAAGT 2640
Qy 2641 ATGAATGTGAAAGGGCATCAACAGCTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTGAAAGGGCATCAACAGCTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
Qy 2701 CTGACGATTATTAAAGAAGTCAATCAGCTCCTGCCAATCAATTTGATAGTGTGCCAGAAATAT 2760
Db 2701 CTGACGATTATTAAAGAAGTCAATCAGCTCCTGCCAATCAATTTGATAGTGTGCCAGAAATAT 2760
Qy 2761 AACACTGAGGATATTGTTTCCCAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCAT 2820
Db 2761 AACACTGAGGATATTGTTTCCCAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCAT 2820
Qy 2821 GTATTTAATAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAAAATTACAATGGACA 2880
Db 2821 GTATTTAATAAACATCGACCCCTGACTTATATGCTTTGGGAACGAAAAATTACAATGGACA 2880
Qy 2881 AATGAACAAATTCAAAGTGCMAAAAGCGGAAATATATCCCGTTTAAACAGTTTCAATAT 2940
Db 2881 AATGAACAAATTCAAAGTGCMAAAAGCGGAAATATATCCCGTTTAAACAGTTTCAATAT 2940
Qy 2941 AATAGTATAAGCTATAAAACATTTTGCATTTTATTAATA 2979
Db 2941 AATAGTATAAGCTATAAAACATTTTGCATTTTATTAATA 2979

RESULT 2

PCT-US99-26501-4

; Sequence 4: Application PC/TUS9926501

; GENERAL INFORMATION:

; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES

; FILE REFERENCE: 617481-5

; CURRENT APPLICATION NUMBER: PCT/US99/26501

; EARLIER FILING DATE: 1999-11-10

; EARLIER FILING DATE: 1998-11-11

; EARLIER FILING DATE: 1999-04-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 2979

; TYPE: DNA

; ORGANISM: Pasteurella multocida

PCT-US99-26501-4

Query Match

Best Local Similarity 100.0%; Score 2979; DB 1: Length 2979;

Matches 2979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATAACTGATTAAAGAGGTTAAAGCTTCAAGCAAGGTTAATTTTAAAGCAAGAAA 60
Db 1 TTATAACTGATTAAAGAGGTTAAAGCTTCAAGCAAGGTTAATTTTAAAGCAAGAAA 60
Qy 61 ATGAATACATTATCAACAGCAATAAAGCATATAACAGCAATGACTTATGAATTAGCACTC 120
Db 61 ATGAATACATTATCAACAGCAATAAAGCATATAACAGCAATGACTTATGAATTAGCACTC 120
Qy 121 AAATTTTGAAGAGTCTGCTGAAACCTACGCGGCAAAATCGTTGAATTCCTCAATATTC 180
Db 121 AAATTTTGAAGAGTCTGCTGAAACCTACGCGGCAAAATCGTTGAATTCCTCAATATTC 180

181 AAATGTAAGAAAAACTCTCGACCAATTCTTATGTAAGTGAAGATAAAAAACAGTGTT 240
181 AAATGTAAGAAAAACTCTCGACCAATTCTTATGTAAGTGAAGATAAAAAACAGTGTT 240
241 TCGGATAGCTCATTAGATATCGCACACAGCTCTTACTTTCCAAACGTAATAAAATAACT 300
241 TCGGATAGCTCATTAGATATCGCACACAGCTCTTACTTTCCAAACGTAATAAAATAACT 300
301 CTATCCGAATCAGAAAAACAGTTTAAAAAAATAATGGAATCTATCACTCGGAAAAAA 360
301 CTATCCGAATCAGAAAAACAGTTTAAAAAAATAATGGAATCTATCACTCGGAAAAAA 360
361 TCGGAGAACGAGAAATCAGAAAGTGGAAGTGGAACTAGTACCACAGATTTTCTTAAGATCTT 420
361 TCGGAGAACGAGAAATCAGAAAGTGGAAGTGGAACTAGTACCACAGATTTTCTTAAGATCTT 420
421 GTTCTTGCTCCATTGGCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA 480
421 GTTCTTGCTCCATTGGCAGATCATGTTAATGATTTTACATGGTACAAAAATCGAAAAAA 480
481 AGCTTAGGTATAAGCCTGTAAATAGAAATATCGGTCCTTCTAATTAATTCCTACATTT 540
481 AGCTTAGGTATAAGCCTGTAAATAGAAATATCGGTCCTTCTAATTAATTCCTACATTT 540
541 AATCGTAGCGGTATTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACCTAC 600
541 AATCGTAGCGGTATTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAAACAACCTAC 600
601 CCATTTGAAGTCGTTGTGCAGATGATGGTGAAGAAAACTTACTTACCATTGTGCAA 660
601 CCATTTGAAGTCGTTGTGCAGATGATGGTGAAGAAAACTTACTTACCATTGTGCAA 660
661 AAATAGCAACAAAACCTTGACATAAAGTATGTAAAGCAAAAAAGATTATGATATCAATTG 720
661 AAATAGCAACAAAACCTTGACATAAAGTATGTAAAGCAAAAAAGATTATGATATCAATTG 720
721 TGTGCACTCAGAACTTAGTGTTCAGTACAGCAAAAGTATGATTTGTCTCGATCTAGAC 780
721 TGTGCACTCAGAACTTAGTGTTCAGTACAGCAAAAGTATGATTTGTCTCGATCTAGAC 780
781 TGGGATATGGCACACCAACAAATATGGGTTCACTTCTTATCTTACAGAACTATTAGAGAC 840
781 TGGGATATGGCACACCAACAAATATGGGTTCACTTCTTATCTTACAGAACTATTAGAGAC 840
841 AATGATATTTGTTAATTTGACCTTAGAAAAATATGTGGATCTCTATAATTTACCGCAGAA 900
841 AATGATATTTGTTAATTTGACCTTAGAAAAATATGTGGATCTCTATAATTTACCGCAGAA 900
901 CAATTCCTTAACGATCCOATATTAATAGAACTACTACTGAAACCGCTACAAATAACAAT 960
901 CAATTCCTTAACGATCCOATATTAATAGAACTACTACTGAAACCGCTACAAATAACAAT 960
961 CTTTCGATTACATCAAAAGAAATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAA 1020
961 CTTTCGATTACATCAAAAGAAATATATCGTTGGATTGGAGATTAGAACATTTCAAAAAA 1020
1021 ACCGATAATCTACGCTCTATGTGATTCCTCGTTTCTGTTATTTGTTGGGGTAATGTTGCA 1080
1021 ACCGATAATCTACGCTCTATGTGATTCCTCGTTTCTGTTATTTGTTGGGGTAATGTTGCA 1080
1081 TTTTCTAAAGAAATGGCTAAATAAAGTATGTTTCCAAAGGCTGTTTTTCCAGAGTAAT 1140
1081 TTTTCTAAAGAAATGGCTAAATAAAGTATGTTTCCAAAGGCTGTTTTTCCAGAGTAAT 1140
1141 GCGGAAGATGTAGAAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTTTTCCAGAGTAAT 1200
1141 GCGGAAGATGTAGAAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTTTTCCAGAGTAAT 1200
1201 GACGGCGAATGGCCATCCATCAAGAACCCCTGGTGAAGAAAAATGAACAGAACCGCAA 1260
1201 GACGGCGAATGGCCATCCATCAAGAACCCCTGGTGAAGAAAAATGAACAGAACCGCAA 1260

1261 GCTGGTAAAAGTATTACGCTTAAAAATTTGCAAGAAAAAGGTACCTTACATCTTATAGAAG 1320
1261 GCTGGTAAAAGTATTACGCTTAAAAATTTGCAAGAAAAAGGTACCTTACATCTTATAGAAG 1320
1321 CTTTACCATAAGAAATTCACATATTCATAGAAATACCTTTAGTTCTTATTTATATCCCC 1380
1321 CTTTACCATAAGAAATTCACATATTCATAGAAATACCTTTAGTTCTTATTTATATCCCC 1380
1381 GCTTATAACTGTCAAAATATATTTCAAGATGTGTAGATAGTCTCTTAAATCAAACTGTT 1440
1381 GCTTATAACTGTCAAAATATATTTCAAGATGTGTAGATAGTCTCTTAAATCAAACTGTT 1440
1441 GTGATCTCGAGTTTGTATTTCTTAACGATGGTTCAACAGATAATACCTTAGAAGTGATC 1500
1441 GTGATCTCGAGTTTGTATTTCTTAACGATGGTTCAACAGATAATACCTTAGAAGTGATC 1500
1501 AATAAGCTTTATCGTAAATAATCTAGGTACGATCATGTCTAAACCAATGCGGGAATA 1560
1501 AATAAGCTTTATCGTAAATAATCTAGGTACGATCATGTCTAAACCAATGCGGGAATA 1560
1561 GCCTCAGCATCAAAATGCAGCGTTTCTTTTGTCTAAAGTTTATTACATTTGGCAGTTAGAT 1620
1561 GCCTCAGCATCAAAATGCAGCGTTTCTTTTGTCTAAAGTTTATTACATTTGGCAGTTAGAT 1620
1621 TCAGATGATTTATCTTGAGCCTGATGCACTGTGTTTAAAGAAATTTTAAAGAT 1680
1621 TCAGATGATTTATCTTGAGCCTGATGCACTGTGTTTAAAGAAATTTTAAAGAT 1680
1681 AAAACGCTACGTTGTGTTTATACCACTAATAGAAACGTCATCCGATGGTAGCTTAATC 1740
1681 AAAACGCTACGTTGTGTTTATACCACTAATAGAAACGTCATCCGATGGTAGCTTAATC 1740
1741 GCTAATGTTTACAAATGGCCAGAAATTTTACAGAGAAAAACTCACAACGGCTATGATTGCT 1800
1741 GCTAATGTTTACAAATGGCCAGAAATTTTACAGAGAAAAACTCACAACGGCTATGATTGCT 1800
1801 CACATTTTGAATGTTTACGATTAGAGCTTGGCATTTTACGGATGGATTTTAAACGAAAT 1860
1801 CACATTTTGAATGTTTACGATTAGAGCTTGGCATTTTACGGATGGATTTTAAACGAAAT 1860
1861 ATTGAAAACCGCGTGGATTTATGACATGTTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
1861 ATTGAAAACCGCGTGGATTTATGACATGTTTAACTCAGTGAAGTTGGAAAAATTTAAA 1920
1921 CATCTTAAATAAATCTGCTATAACCGGCTATTACATGGTGATACACATCCATTAAAGAA 1980
1921 CATCTTAAATAAATCTGCTATAACCGGCTATTACATGGTGATACACATCCATTAAAGAA 1980
1981 CTCGGCATTTCAAAAGAAAAACCATTTTCTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
1981 CTCGGCATTTCAAAAGAAAAACCATTTTCTGTAGTCAATCAGTCAATTAATAGACAAGGC 2040
2041 ATCAATTTATTAATTTATGCAAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTC 2100
2041 ATCAATTTATTAATTTATGCAAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTC 2100
2101 AATAAAACCGCTGAATATCAAGAGAAATGATATGTTTAAAGATCTTAACTCATTCAA 2160
2101 AATAAAACCGCTGAATATCAAGAGAAATGATATGTTTAAAGATCTTAACTCATTCAA 2160
2161 AATAAAGATGCAAAATCGCAGTCAGTATTTTCTATCCCAATACATTAACCGCTTAGTG 2220
2161 AATAAAGATGCAAAATCGCAGTCAGTATTTTCTATCCCAATACATTAACCGCTTAGTG 2220
2221 AAAAACTTAAACAATTTATTTGAATATAATAAATAATATTCGTTTATTTTACATGTT 2280
2221 AAAAACTTAAACAATTTATTTGAATATAATAAATAATATTCGTTTATTTTACATGTT 2280
2281 GATAGAAATCATCTTTACACGACATCAAAAAGAAATTTGCGTTTCTTATCATAGCAC 2340
2281 GATAGAAATCATCTTTACACGACATCAAAAAGAAATTTGCGTTTCTTATCATAGCAC 2340
2341 CAAGTGAATATTTTACTAAATAATGACATCTCATATTACACGAGTAATAGACTAATAA 2400

QY 1141 GCGAAGATGTAGAAATTTGGTTACAGATTATTTGCCAAGGCGTGTCTTTTTCAGAGTAATT 1200
DB 1141 GCGAAGATGTAGAAATTTGGTTACAGATTATTTGCCAAGGCGTGTCTTTTTCAGAGTAATT 1200
QY 1201 GACGGCGGAATGCCATCCATCAAGAACCCACCTGGTTAAAGAAAATGAACAGAACCGGAA 1260
DB 1201 GACGGCGGAATGCCATCCATCAAGAACCCACCTGGTTAAAGAAAATGAACAGAACCGGAA 1260
QY 1261 GCTGGTAAAGTATACGCTTAAATTTGAAAGAAAGGTACCTTACATCTATAGAAAG 1320
DB 1261 GCTGGTAAAGTATACGCTTAAATTTGAAAGAAAGGTACCTTACATCTATAGAAAG 1320
QY 1321 CTTTAAACCAATGAAGATTCACATATTCATAGAAATACCTTTAGTTCTTATATATACCC 1380
DB 1321 CTTTAAACCAATGAAGATTCACATATTCATAGAAATACCTTTAGTTCTTATATATACCC 1380
QY 1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTGTAGATAGTCTCTTAATCAAACTGTT 1440
DB 1381 GCTTATAACTGTGCAAAATATATTCAAAGATGTGTAGATAGTCTCTTAATCAAACTGTT 1440
QY 1441 GTCGATCTCGAGGTTGTTATTTGTAACGATGGTTCAACAGATAATACCTTAGAAGTGATC 1500
DB 1441 GTCGATCTCGAGGTTGTTATTTGTAACGATGGTTCAACAGATAATACCTTAGAAGTGATC 1500
QY 1501 AATAAGCTTTATGGTAAATATCTAGGTACGATCATGTCTAAACCAATGGCGGAATA 1560
DB 1501 AATAAGCTTTATGGTAAATATCTAGGTACGATCATGTCTAAACCAATGGCGGAATA 1560
QY 1561 GCCTCAGATCAAAATGCAGCGGTTCTTTTGCCTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
DB 1561 GCCTCAGATCAAAATGCAGCGGTTCTTTTGCCTAAAGGTTATTACATTTGGCGAGTTAGAT 1620
QY 1621 TCAGATGATTATCTTGAGCCTGATGAGTTGAACCTGTTTAAAGAAATTTTAAAGAGAT 1680
DB 1621 TCAGATGATTATCTTGAGCCTGATGAGTTGAACCTGTTTAAAGAAATTTTAAAGAGAT 1680
QY 1681 AAAACGCTAGCTGTGTTTATACCACTTAATAGAAACGTCATCCGATGGTGTAGCTTAATC 1740
DB 1681 AAAACGCTAGCTGTGTTTATACCACTTAATAGAAACGTCATCCGATGGTGTAGCTTAATC 1740
QY 1741 GCTAATGTTTACAAATGGCCAGATTTTACAGAGAAACCTCAACGCGCTATGATGCT 1800
DB 1741 GCTAATGTTTACAAATGGCCAGATTTTACAGAGAAACCTCAACGCGCTATGATGCT 1800
QY 1801 CACCAATTTTGAATTTTACGATTTAGAGCTTGGCAATTTAAACGATGGATTTAAACGAAAT 1860
DB 1801 CACCAATTTTGAATTTTACGATTTAGAGCTTGGCAATTTAAACGATGGATTTAAACGAAAT 1860
QY 1861 ATTGAAACGCGGTGATTTAGCATGTTCCCTTAACTCAGTGAAGTTGGAATAATTTAAA 1920
DB 1861 ATTGAAACGCGGTGATTTAGCATGTTCCCTTAACTCAGTGAAGTTGGAATAATTTAAA 1920
QY 1921 CATCTTAATAAATCTGCTATAACCGGTTATACATGGTGATAACACATCCATTAAAGAA 1980
DB 1921 CATCTTAATAAATCTGCTATAACCGGTTATACATGGTGATAACACATCCATTAAAGAA 1980
QY 1981 CTCGGCATTTCAAAAGAAAACCAATTTGTTGTAGTCAATCAGTCAATTAATAGCAAGGC 2040
DB 1981 CTCGGCATTTCAAAAGAAAACCAATTTGTTGTAGTCAATCAGTCAATTAATAGCAAGGC 2040
QY 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTTC 2100
DB 2041 ATCAATTTATTAATATGACAAATTTGATGATTTAGATGAAGTAGAAGTATATCTTTC 2100
QY 2101 AATAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
DB 2101 AATAAACCGCTGAATATCAAGAGAAATGGATATGTTTAAAGATCTTAAACTCATTTCAA 2160
QY 2161 AATAAAGATGCAAAATGCAGTCAAGTATTTCTATCCCAATACATTAACGCGCTTAGTG 2220
DB 2161 AATAAAGATGCAAAATGCAGTCAAGTATTTCTATCCCAATACATTAACGCGCTTAGTG 2220
QY 2221 AAAAACTTAACAATATTTATTTGAATATAATAAAAAATATATTCGTTATTATTCTACATGTT 2280

DB 2221 AAAAACTTAACAATATTTATTTGAATATAATAAAAAATATATTCGTTATTATTCTACATGTT 2280
QY 2281 GATAAGAATCATCTTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAGCAC 2340
DB 2281 GATAAGAATCATCTTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAAGCAC 2340
QY 2341 CAAAGTGAATATTTTACTTAAATATGACATCTCATATTTACACGAGTAATAGACTAATAAAA 2400
DB 2341 CAAAGTGAATATTTTACTTAAATATGACATCTCATATTTACACGAGTAATAGACTAATAAAA 2400
QY 2401 ACTGAGGCACATTTTAGTAAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTAATAC 2460
DB 2401 ACTGAGGCACATTTTAGTAAATATTAATAATTAAGTCAGTTAAATCTAAATTTGTAATAC 2460
QY 2461 ATCATTTTTCATATCATGACAGCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
DB 2461 ATCATTTTTCATATCATGACAGCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
QY 2521 AAATATGATCTCGCATGAATTTTCTCAGCATTAACACATGATTTGATCGAGAAAAATCAAT 2580
DB 2521 AAATATGATCTCGCATGAATTTTCTCAGCATTAACACATGATTTGATCGAGAAAAATCAAT 2580
QY 2581 GCGCATCCACATTTTAAAGAGCTGAATTTAAACCTATTTTAAATGACATGACTTAAAGAGT 2640
DB 2581 GCGCATCCACATTTTAAAGAGCTGAATTTAAACCTATTTTAAATGACATGACTTAAAGAGT 2640
QY 2641 ATGAATGTGAAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
DB 2641 ATGAATGTGAAAGGGGCATCACAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT 2700
QY 2701 CTGACGATTTTAAAGAGCTATCACATCTCGCAATCAATTTGATGTCGCCAGATAT 2760
DB 2701 CTGACGATTTTAAAGAGCTATCACATCTCGCAATCAATTTGATGTCGCCAGATAT 2760
QY 2761 AACACTGAGGATATTTGTTCCCAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCAT 2820
DB 2761 AACACTGAGGATATTTGTTCCCAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCAT 2820
QY 2821 GTATTTTAAACATCGACCTGACTTATATGCTTTGGGAACGAAAAATTTCAATTTGAGCA 2880
DB 2821 GTATTTTAAACATCGACCTGACTTATATGCTTTGGGAACGAAAAATTTCAATTTGAGCA 2880
QY 2881 AATCAACAAATTTCAAGTGCAAAAAGGCGAAAAATATCCCGTTTAAACAAGTTCAATT 2940
DB 2881 AATCAACAAATTTCAAGTGCAAAAAGGCGAAAAATATCCCGTTTAAACAAGTTCAATT 2940
QY 2941 AATAGTATAACGCTATAAACAATTTGCAATTTTATTTAAAA 2979
DB 2941 AATAGTATAACGCTATAAACAATTTGCAATTTTATTTAAAA 2979

RESULT 4
US-10-011-768B-8
; Sequence 8, Application US/10011768B
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011,768B
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: Pastuerella Multocida

US-10-011-768B-8

Query Match		76.7%;	Score 2285;	DB 38;	Length 2937;
Best Local Similarity		86.5%;	Pred. No. 0;		
Matches 2541;		Conservative	0;	Mismatches	375;
				Indels	21;
				Gaps	1;
Qy	43	ATTTTAAAGGAAGAAATGAATACATTTATCAACGCAATATAAAGCATATAACAGCAAT	102		
Db	1	ATTTTAAAGGACAGAAATGAATACATTTATCAACGCAATATAAAGCATATAACAGCAAT	60		
Qy	103	GACTATGAATAGCAGCTCAATATTTTGAAGAGTCTGCTGAAACCTAGCGCGGAAATATC	162		
Db	61	GACTATCAATAGCAGCTCAATATTTTGAAGAGTCTGCGGAAATCTATGGACGGAATTT	120		
Qy	163	GTGGAATTCGAATTTATCAATGTAAAGAAAACCTCTCGACCAATTC-----	209		
Db	121	GTGGAATTTCAATTTACCAATGCAAGAAAACCTCTCAGCACATCTCTGTTAAATTC	180		
Qy	210	-----TTATGTAAGTGAAGATAAAGAAAACAGTGTGTCGATAGCTCATATGATATC	261		
Db	181	GCACATCTTCTGTAATAAAGAAAGAAAAGTCAATGTTGCGATAGTCGCTTAGATATT	240		
Qy	262	GCAACAGCTCTTACTTCCAGCTGAAAATAATTAACCTATCCGCAATCAGAAAAC	321		
Db	241	GCAACAGCTCTTACTTCCAGCTGAAAATAATTAACCTATCCGCAATCAGAAAAC	300		
Qy	322	AGTTTAAATAAATAATGAAATCTATCACCTGGGAAAATAATCGGAGAACGCAAGAAATC	381		
Db	301	AGTTTAAATAAATAATGAAATCTCTCTGAGAGAAATCTGAAATGCGGAGGTAA	360		
Qy	382	AAAGTGAAGTACGACCAAGATTTTCCCTAAAGATTTGTTCTGCTCCATTTCCAGAT	441		
Db	361	CGGCTCGCCCTTGTACCAAAAGATTTTCCCAAGATCTGTTTACGCGCTTTACCTGAT	420		
Qy	442	CATGTTAATGATTTACATGTCACAAATTCGAAAATAAGCTTAGGTATPAAAGCCTGTA	501		
Db	421	CATGTTAATGATTTACATGTCACAAATTCGAAAATAAGCTTAGGTATPAAAGCCTGTA	480		
Qy	502	AAAGAAATATCGGCTCTCTATTTATTTCTCATATTTAATCTGAGCGGTATTTAGAT	561		
Db	481	CATCAACATGTTGGCTCTCTATTTATTTCTGTTACACATTTCAATCGACAGCAATTTAT	540		
Qy	562	ATACGTTAGCTGTTGGTCAATCAGAAAACAACTACCCATTTGAAAGTCTGTTGCTGCA	621		
Db	541	ATTTACATTTAGCTGTTTGTAGTAAACCAACAAACACATTTACCGCTTTGAGTTATCGTG	600		
Qy	622	GATGATGCTAGTAAAGAAATCTACTTACCAATGTCGAAAATAACGACAAAACCTGAC	681		
Db	601	GATGATGCTAGTAAAGAAATCTACTTACCGGATCATTCGCCAATATGAAAATAAATTTG	660		
Qy	682	ATAAGATATGTAACACAAAAGATTTATGGATATCAATTTGTCGAGTCAGAAAACCTTAG	741		
Db	661	ATTCGCTACGTCAGACAAAAGATTAACGGTTTCAAGCCAGTGGCTCGGAATATGGGA	720		
Qy	742	TTACGTACAGCAAAATGATTTTGTCTCGATTTCTAGACTCGGATATGGCACCACAA	801		
Db	721	TTACGCTTAGCAAAATATGACTTTTATTGGCTTACTCGACTGTGATATGGCGCCAAATCC	780		
Qy	802	TTATGGGTTCAATCTTATCTACAGAACTATTAGAAGCAATGATATGTTTAAATGGA	861		
Db	781	TTATGGGTTCAATCTTATGTTGACAGCTATTAGAAGATGATGATTTAACAATCAATCTG	840		
Qy	862	CCTAGAAAATATGGATACATCAATATTTACCGCAGAACAAATTCCTTTAACGATCCAT	921		
Db	841	CCAAGAAAATACATCGATACACACATATTTGACCCAAAAGACTTCTTAAATTAACCGAG	900		
Qy	922	TTAATAGAATCACTACCTGAAACCGCTACAAATAACAAATCCTCGATTTACATCAAAAG	981		
Db	901	TTGCTTGAATCATTTACCAGAAAGTGAACCAATAATAGTGTGCGGCAAAAGGGGAAG	960		
Qy	982	AAATATATCGTTGGATTTAGACATTTTCAAAAATAACCGGATAATCTACGCTTATGT	1041		
Db	961	ACGTTTCTGCTGATTTGGCGCTTAGAACAAATTCGAAAAAACAAGAAATCTCCGCTTATCC	1020		

Qy	1042	GATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTTGCATTTTCTAAAGAAATGGCTAAAT	1101		
Db	1021	GATTCGCTTTTCGTTTTCGGCGGGTAAATGTTGCTTTCGCTTAAAGAAATGGCTAAAT	1080		
Qy	1102	AAAGTAGTGTGTTGCGATGAAGAATTTAATCATTTGGGGGGCGAAGATGTAAGATTTGGT	1161		
Db	1081	AAATCCGGTTCCTTTGATGAGGAATTTAATCACTGGGTGGAGAAGATGTGAATTTGGA	1140		
Qy	1162	TACAGATTTATTTGCCAAAGGCTGTTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCAT	1221		
Db	1141	TATCGCTTTATTTCCGTTACGGTAGTTTCTTTAAACATATTTGATGGCATTTATGCCCTACC	1200		
Qy	1222	CAAGAACCCCTTGGTAAAGAAAAATGAAACAGAACGCGAAGCTGGTAAAAAGTATTACGGCT	1281		
Db	1201	CAAGACCCACCGATTAAAGAAAAATGAACCGGATCGTGAACGGGAAAGAAATATTACGCTC	1260		
Qy	1282	AAAATTGTGAAGAAAAGTACCTTTACATCTATAGAAGCTTTTACCAATAGAAGATTC	1341		
Db	1261	GATATTATGAGAAAAGTCCCTTATATCTATAGAAAACTTTACCAATAGAAGATTCG	1320		
Qy	1342	CATATTTCATAGATACCTTTAGTTTCTATTTATATCCCGCTTATAACTGTGCAATTTAT	1401		
Db	1321	CATATCAATAGATACCTTTAGTTTCAATTTATATCCCGCTTATAACTGTGCAAACTAT	1380		
Qy	1402	ATTCAAAGATGTGTAGATAGTCTCTTAATCAAACTGTTGTCTGATCTCCAGGTTTGTATT	1461		
Db	1381	ATTCACCGTTGCGTAGATAGTGCACCTGAATCAGACTGTTGTTGATCTCCAGGTTTGTATT	1440		
Qy	1462	TGTAACGATGTTCAACAGATAAATACCTTTAGAAGTGATCAATAAGCTTTTATGGTAAAT	1521		
Db	1441	TGTAACGATGTTCAACAGATAAATACCTTTAGAAGTGATCAATAAGCTTTTATGGTAAAT	1500		
Qy	1522	CCTAGGTTACGATCATGTCTAAACCAATGCGGAATAGCCTCAGCATCAATATGACGC	1581		
Db	1501	CCTAGGTTACGATCATGTCTAAACCAATGCGGAATAGCCTCAGCATCAATATGACGC	1560		
Qy	1582	GTTCTCTTTTGTAAAGGTTATTACATTTGGCAGTTTAGATTTCAGATGATTATCTTGAGCCT	1641		
Db	1561	GTTCTCTTTTGTAAAGGTTATTACATTTGGCAGTTTAGATTTCAGATGATTATCTTGAGCCT	1620		
Qy	1642	GATGAGTTGAACCTGTGTTTAAAGAAATTTTAAAGATPAAACGCTAGCTTGTGTTTAT	1701		
Db	1621	GATGAGTTGAACCTGTGTTTAAAGAAATTTTAAAGATPAAACGCTAGCTTGTGTTTAT	1680		
Qy	1702	ACCACTAATAGAAACGTCATCCCGATGCTAGCTTTAATCGCTTAATGGTTTACAAATGGCCA	1761		
Db	1681	ACCACTAATAGAAACGTCATCCCGATGCTAGCTTTAATCGCTTAATGGTTTACAAATGGCCA	1740		
Qy	1762	GAATTTTTCAGGAGAAAACCTCACACGGCTATGATTGCTCAGCATTTTAGAATGTTTAGC	1821		
Db	1741	GAATTTTTCAGGAGAAAACCTCACACGGCTATGATTGCTCAGCATTTTAGAATGTTTAGC	1800		
Qy	1822	ATTAGAGCTTGGCATTTTAAACGGATGATTTAACGAAAAATATTGAAAACCGCGGTGATAT	1881		
Db	1801	ATTAGAGCTTGGCATTTTAACTGATGATTTCAATGAAAAAATTTGAAAAATCCGCTAGACT	1860		
Qy	1882	GACATGTTCCCTTAAACTCAGTCAAGCTGGAATAATTTAAACATCTTTAATAAATCTGCTAT	1941		
Db	1861	GACATGTTCCCTTAAACTCAGTCAAGCTGGAATAATTTAAACATCTTTAATAAATCTGCTAT	1920		
Qy	1942	AACCCGCTATTACATGTTGATTAACACATCCATTAAGAAAACCTGGCATTTCAAAGAAAAAC	2001		
Db	1921	AACCCGCTATTACATGTTGATTAACACATCCATTAAGAAAACCTGGCATTTCAAAGAAAAAC	1980		
Qy	2002	CATTTTGTGTAGTCAATCAGTCATTTAAATAGACAGGATCAATTTATTTATTTATGAC	2061		
Db	1981	CATTTTGTGTAGTCAATCAGTCATTTAAATAGACAGGATCAATTTATTTATTTATGAC	2040		
Qy	2062	AAATTTGATGATTTAGATGAAAAGTAGAAGTATATCTCAATAAAACCCCTGAATATCAA	2121		
Db	2041	GAATTTGATGATTTAGATGAAAAGTAGAAGTATATTTTCAATAAAACCCCTGAATATCAA	2100		

Qy	2122	GAAGAAATGGATATCTTTAAAGATCTTAAACTATTCAAAATAAAGATGCCAAAATCGCA	2181
Db	2101	CAAGAGATTGATCTTTAAAGATATTAAATCATCCAGATAAAGATGCCAAAATCGCA	2160
Qy	2182	GTCACTATTCTTATCCCAATACATTAACGGCTTAGTGAAATAACTAAACAATATTAT	2241
Db	2161	GTCACTATTCTTATCCCAATACATTAACGGCTTAGTGAAATAACTAAACAATATTAT	2220
Qy	2242	GAATATAATAAAATATATCTGTTATTATCTACATGATGATGATGATGATGATGATGAT	2301
Db	2221	GAATATAATAAAATATATCTGTTATTATCTACATGATGATGATGATGATGATGATGAT	2280
Qy	2302	GACATCAAAAAGAAATATTGCTTTCTATCATATAGCACCAGTGAATATTTTACTATAT	2361
Db	2281	GATATCAAAAAGAAATATTGCTTTCTATCATATAGCACCAGTGAATATTTTACTATAT	2340
Qy	2362	AATGACATCTCATATTACACGAGTAAATAGACTAATAAATACTGAGGCACATTTAAGTAT	2421
Db	2341	AATGATATCTCATATTACACGAGTAAATAGACTAATAAATACTGAGGCACATTTAAGTAT	2400
Qy	2422	ATTAATAAATTAAGTCAGTTAAATCTAAATGTGTAATACATCTTTTGTATATCATGAC	2481
Db	2401	ATTAATAAATTAAGTCAGTTAAATCTAAATGTGTAATACATCTTTTGTATATCATGAC	2460
Qy	2482	AGCCTATTGTTTAAATGACAGCTATGCTTATATGAAAAATATATGATGCGCATGAAT	2541
Db	2461	AGCCTATTGTTTAAATGACAGCTATGCTTATATGAAAAATATATGATGCGCATGAAT	2520
Qy	2542	TTCTCAGCATTAACACATGATTTGATCGAGAAATCAATGCGCATCCACCATTTAAAG	2601
Db	2521	TTCTCAGCATTAACACATGATTTGATCGAGAAATCAATGCGCATCCACCATTTAAAG	2580
Qy	2602	CTCATTAAACCTATTATTAATGACATGACTTAAGAGTATGAATGTGAAGGGGCATCA	2661
Db	2581	CTCATTAAACCTATTATTAATGACATGACTTAAGAGTATGAATGTGAAGGGGCATCA	2640
Qy	2662	CAAGGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTAATAAAGATC	2721
Db	2641	CAAGGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTAATAAAGATC	2700
Qy	2722	ATCACATCTGCAATCAATGATGATGTCGAGAAATATAACATGAGGATATTGTTTC	2781
Db	2701	ATCACATCTGCAATCAATGATGATGTCGAGAAATATAACATGAGGATATTGTTTC	2760
Qy	2782	CAATTTGCACTTTAATCTTAGAAAAGAAACCGGCATGATTTAATAAATCAATCGACC	2841
Db	2761	CAATTTGCACTTTAATCTTAGAAAAGAAACCGGCATGATTTAATAAATCAATCGACC	2820
Qy	2842	CTGACTTATATGCTTTGGGACGAAATTAACATGGACAAATGAACAAATTCAAAGTGA	2901
Db	2821	CTGACTTATATGCTTTGGGACGAAATTAACATGGACAAATGAACAAATTCAAAGTGA	2880
Qy	2902	AAAAAGGGGAAATATCCCGTTAAACAGTTCATTTAATAGTATAACGCTATAA	2958
Db	2881	AAAAAGGGGAAATATACCTGTTTAAACAGTTCATTTAATAGTATAACCTATAA	2937
RESULT 5			
US-10-011-771B-8			
; Sequence 8, Application US/10011771B			
; GENERAL INFORMATION: Paul H.			
; APPLICANT: Weigel, Paul			
; APPLICANT: DeAngelis, Paul			
; APPLICANT: Kumari, Kshama			
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof			
; FILE REFERENCE: 3554.011			
; CURRENT APPLICATION NUMBER: US/10/011.771B			
; CURRENT FILING DATE: 2001-10-11			
; PRIOR APPLICATION NUMBER: US 09/178,851			
; PRIOR FILING DATE: 1998-10-26			
; PRIOR APPLICATION NUMBER: US 60/064,435			
; PRIOR FILING DATE: 1997-10-31			
; NUMBER OF SEQ ID NOS: 10			

;
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: Pastuerella Multocida
US-10-011-771B-8

Query Match

Best Local Similarity

Matches 2541

76.7%;

Score 2285;

DB 38;

Length 2937;

86.5%;

Pred. No. 0;

0;

Mismatches 375;

Indels 21;

Gaps 1;

Qy	43	ATTATTAAAGGAAAGAAATGAATACATTATCACAGCAATAAAGCATATACACGAAT	102
Db	1	ATTATTAAAGGACAGAAATGAATACATTATCACAGCAATAAAGCATATACACGAAT	60
Qy	103	GACTATGAATTAGCACTCAAAATTTTGGAACTGCTGAAACCTACGGCGGAAATATC	162
Db	61	GACTATCAATTAGCACTCAAAATTTTGGAACTGCTGAAACCTACGGCGGAAATATC	120
Qy	163	GTTGAATTCCAAATTTATCAAAATGTAAGAAAAAACTCTCGACCAATTC	209
Db	121	GTTGAATTTCAAATTTACCAATGCAAAAGAAAACTCTCAGCACATCTCTGTTAATTC	180
Qy	210	-----TTATGTAAGTGAAGATAAAGAAAAACAGTCTTTCGATAGCTCATTAGATATC	261
Db	181	GCACATCTTTCTGTAATAAAGAAAAAGTCAATGTTTCGATAGCTCATTAGATATC	240
Qy	262	GCAACACAGCTCTTACTTTCCAACTGTAAGAAAAATTAATCTCTATCCGAATCAGAAAAAAC	321
Db	241	GCAACACAGCTCTTACTTTCCAACTGTAAGAAAAATTAATCTCTGACTCGGAAAAAAC	300
Qy	322	AGTTTAAAAAATAAATGAAATCTATCAGTGGGAAAAAATCGGAGAACGCAAGATCAGA	381
Db	301	ACGTTAAAAAATAAATGAAATTCCTCAGTGAAGAAAAATCTGAAATCGGAGGTAGA	360
Qy	382	AAAGTGAACCTAGTACCAAGATTTTCCCTAAAGATCTGTTCTGCTCCATTCGCCAGAT	441
Db	361	CGGTCGCCCTTGTAACCAAGATTTTCCCAAGATCTGTTTTCGCTTACCTGAT	420
Qy	442	CATGTTAATGATTTTACATGGTACAAAAATCGAAAAAAGCTTAGGTATTAAGCCTGTA	501
Db	421	CATGTTAATGATTTTACATGGTACAAAAAAGCAAGAAAGCTTGGCATAAAACCTGAA	480
Qy	502	AATGAAGATATCGCTCTCTTCTATTATTTCTCTATTAATCTGATGAGCGGTATTTAGAT	561
Db	481	CATCAACATGTTGCTCTCTTCTATTATCGTTACAACTCAATCGACCAATTTATATCG	540
Qy	562	ATAAGCTTAGCTGTTGGTCAATCAGAAAAAACAACCTACCCATTTGAAGTCGTTGCA	621
Db	541	ATTACATTAGCTGTTTAGTAAACCAAAACACATTAACCGTTTGAAGTTATCGTGACA	600
Qy	622	GATGATGATGTAAGAAAACTTACTTACCATTGTGCAAAATACGAAACAAACCTTAC	681
Db	601	GATGATGATGTAAGAAAACTTACTTACCATTGTGCAAAATACGAAACAAACCTTAC	660
Qy	682	ATAAGATGTAAGAAAAAGATTTATCGATCAATTTGTGTCAGTGCAGTGCAGTGCAG	741
Db	661	ATTCGCTACGTCAGAAAAAGATTAACCGTTTCAAGCCAGTGCCTCGGAATATGGA	720
Qy	742	TTAGTACAGCAAAAGTATGATTTTCTCGATCTTAGCTGCGATGCGGATGCGCACCACA	801
Db	721	TTAGCTTTAGCAAAATATGACTTTTATGGCTTACTCGACTGTGATATGCGGCAATCA	780
Qy	802	TTATGGTTTCAATCTTATCTTACAGAACTATTAGAGACAAATGATTTCTTTAATTTGA	861
Db	781	TTATGGTTTCAATCTTATGTCAGAGCTATTAGAGATGATTTTAAACATCATTTGTT	840
Qy	862	CCTAGAAAAATGTCGATACTCATATAATTAATCCGAGACAAATCTCTTAACGATCATAT	921
Db	841	CCAAGAAAAATGATCGATACACAACTATTGACCCCAAGAGCTTCTTAAATAACGGAGT	900
Qy	922	TTAATAGAACTACTACCTGAAACCGGTACAAATAACAATCTCTCGATTACATCAAGGA	981

Db 901 TTGCTTGAATCATTACCAGAAAGTGAAACCAATAAATAGTGTGGCGCAAAAGGGGAAGGA 960
Qy 982 AATATATCGTTGGATTGAGATTTAGAACATTTCAAAACAAACCGATAATCTACGCTATAGT 1041
Db 961 ACAGTTTCTCGGATTTGGCGCTTTAGAACAAATTCGAAACAAACAGAAAATCTCCGCTTATCC 1020
Qy 1042 GATTCTCCGTTTCGTTATTTTGTGGGGTAATGTTGCGATTTTCTAAAGAAATGGCTAAAT 1101
Db 1021 GATTGCGCTTTCCGCTTTTGTGGCGGGTAATGTTGCTTTCGCTAAAATGGCTAAAT 1080
Qy 1102 AAAGTAGTGGTTTCGATGAAGAATTTAATCATTTGGGGGGGGGAGAGTGAATTTGGT 1161
Db 1081 AAATCCGGTTTCTTTGTATGAGGAATTTAATCACTGGGGTGAGAAGAGTGTGGAAATTTGGA 1140
Qy 1162 TACAGATTATTTGCAAGGCTGTTTTTTCAGAGTAATTTGACGGGGAATGGCCATCCAT 1221
Db 1141 TATCGCTTATTTCCGTTACGGTAGTTTCTTTTAAACATATTGATGGCATTTATGCCCTACCAT 1200
Qy 1222 CAAGAACCCACTGGTTAAAGAAAATGAACAGAACGCGAAGCTGGTAAAAGTATTTACGCTT 1281
Db 1201 CAAGAGCCACCAGGTAAAGAAAATGAACCCGATCGTGAAGCGGGAANAATATTACGCTC 1260
Qy 1282 AAAATTTGGAAGAAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAAGATTCA 1341
Db 1261 GATATTATGAGAGAAAGGTCCCTTATATCTATAGAAAACCTTTTACCAATAGAAGATTCCG 1320
Qy 1342 CATATTATAGAAATACCTTTAGTTCTATTTTATATCCCGCTTATTAAGTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTTATGTTTCAATTTATATCCAGCTTATTAAGTGTGCAAAATAT 1380
Qy 1402 ATTCAAAAGATGTAGATAGTGCCTTTAATCAAACTGTGTCGATCTCGAGGTTTGTATT 1461
Db 1381 ATTCACAGTTTCGTTAGATAGTGCCTGAATCAGACTGTTGTTGATCTCGAGGTTGTATT 1440
Qy 1462 TGTAAAGATGTTCAACAGATAATACCTTAGAAGTGAATCAATAGCTTTTATGGTAATAT 1521
Db 1441 TGTAAAGATGTTTCAACAGATAATACCTTAGAAGTGAATCAATAGCTTTTATGGTAATAT 1500
Qy 1522 CCTAGGTTACGCATCATCTCTAAACCAATGCGGAATAGCTTCAGCATCAAAATCGACCC 1581
Db 1501 CCTAGGTTACGCATCATCTCTAAACCAATGCGGAATAGCTTCAGCATCAAAATCGACCC 1560
Qy 1582 GTTCTTTTGTAAAGGTTATTACATTTGGGCGAGTTAGATTTCAGATGATTATCTTGAGGCT 1641
Db 1561 GTTCTTTTGTAAAGGTTATTACATTTGGGCGAGTTAGATTTCAGATGATTATCTTGAGGCT 1620
Qy 1642 GATCGAGTTGAACGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTTGTGTTTAT 1701
Db 1621 GATCGAGTTGAACGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTTGTGTTTAT 1680
Qy 1702 ACCACTAATAGAAACGTTCAATCCGATGGTAGCTTAATCGCTAATGGTTACAAATGGGCA 1761
Db 1681 ACCACTAATAGAAACGTTCAATCCGATGGTAGCTTAATCGCTAATGGTTACAAATGGGCA 1740
Qy 1762 GAATTTTTCACGAGAAAACCTCACAAACGCTATGATGCTCACCATTTTAGAAATGTTTACG 1821
Db 1741 GAATTTTTCACGAGAAAACCTCACAAACGCTATGATGCTCACCATTTTAGAAATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTAACGGATGGATTAAACGAAATATTGAAACCGCGTGATTTAT 1881
Db 1801 ATTAGAGCTTGGCATTTAACCTGATGGATTCAATGAAAAAATTTGAAATCCCGCTAGACTAT 1860
Qy 1882 GACATGTTCTTAACTACAGTGAAGTTGGAATAATTTAAACATCTTAATAAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTACAGTGAAGTTGGAATAATTTAAACATCTTAATAAATCTGCTAT 1920
Qy 1942 AACCGCGTATTACATGTTGATGAACACATCTTAAAGAAACCTCGGATTTCAAAAGAAAAC 2001
Db 1921 AACCGGTGATTACATGTTGATGAACACATCAATTAAGAACTTGGCATTTCAAGAAAC 1980
Qy 2002 CATTTTGTGTAGTCAATCACTCATTAATAGAACGGCATCAATTTATTAATTTATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCACTCATTAATAGAACGGCATCAATTTATTAATTTATGAC 2040

Qy 2062 AAATTTGATGATTTAGATGAAAGTAGAAAGTATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAAGTAGAAAGTATATTTTCAATAAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATATGTTTAAAGATCTTAAACATCTTCAAAATTAAGATGCCAAAATCGCA 2181
Db 2101 GAAGAGATTGATATCTTAAAGATATTTAAATCATCCAGAAATAAAGATGCCAAAATCGCA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACAATATATT 2241
Db 2161 GTCAGTATTTTCTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACAATATATT 2220
Qy 2242 GAATTAATAAATAATATTTCTGTTTATTTCTACATGTTGATAAGAAATCATCTTACACCA 2301
Db 2221 GAATTAATAAATAATATTTCTGTTTATTTCTACATGTTGATAAGAAATCATCTTACACCA 2280
Qy 2302 GACATCAAAAAGAAATATTTGGCTTTCTATCATCAAGCACCAAGTAGTAATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATACTAGCTTCTATCATAAACATCAAGTGAATATTTTACTAAAT 2340
Qy 2362 AATGACATCTCATTTTACAGAGTAAATGACACTAAATAAAGCTGAGGACATTTTAAAGTAAT 2421
Db 2341 AATGATATCTCATTTTACAGAGTAAATGACACTAAATAAAGCTGAGGACATTTTAAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTTAAATCTTAAATCTGTAATACATCATTTTGTATATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTTAAATCTGTAATACATCATTTTGTATATCATGAC 2460
Qy 2482 AGCCTATTCGTTTAAAGATGACAGTATGCTTATATGAAAAAATATATGTCGCGCATGAAT 2541
Db 2461 AGCCTATTCGTTTAAAGATGACAGTATGCTTATATGAAAAAATATGTCGCGCATGAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAATGCGCATCCACCATTTTAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTGGATCGAGAAAAATCAATGCGCATCCACCATTTTAAAG 2580
Qy 2602 CTGATTAACACCTATTTTAAATGACATGACTTAAAGAAATGATGAATGTGAAAGGGGATCA 2661
Db 2581 CTCATTAACACCTATTTTAAATGACATGACTTAAAGAAATGATGAATGTGAAAGGGGATCA 2640
Qy 2662 CAAGTATGTTTATGAAGTATGCGTACCGCATGAGCTTCTGACGATTTTAAAGAAAGTC 2721
Db 2641 CAAGTATGTTTATGACGATGCGTACCGCATGAGCTTCTGACGATTTTAAAGAAAGTC 2700
Qy 2722 ATCAATCTCGCAATCAATTTGATAGTGTGCGAGAAATATAACACTGAGGATATTTGGTTC 2781
Db 2701 ATCAATCTCGCAATCAATTTGATAGTGTGCGAGAAATATAACACTGAGGATATTTGGTTC 2760
Qy 2782 CAATTTGCACTTTTAACTTTAGAAAAAGAAACCGGCGCATGTTTAAATAAACATCGACC 2841
Db 2761 CAATTTGCACTTTTAACTTTAGAAAAAGAAACCGGCGCATGTTTAAATAAACATCGACC 2820
Qy 2842 CTGACTTATATGCTTTGGGAACGAAATTTACAATGGACAAATGAACAAATTTCAAAGTGCA 2901
Db 2821 CTGACTTATATGCTTTGGGAACGAAATTTACAATGGACAAATGAACAAATTTGAAAGTGCA 2880
Qy 2902 AAAAAAGGCGAAATATCCCGTTTAAACAAGTTTCATTTAATAAGTATACCGCTATRA 2958
Db 2881 AAAAAAGGCGAAATATACCTGTTTAAACAAGTTTCATTTAATAAGTATACCTCTATRA 2937

RESULT 6

PCT-US99-07289-2

; Sequence 2, Application PC/TUS9907289

; GENERAL INFORMATION:

; APPLICANT: Board of Regents of the University of Oklahoma

; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS

; TITLE OF INVENTION: OF USE

; FILE REFERENCE: 617022-7

; CURRENT APPLICATION NUMBER: PCT/US99/07289

; CURRENT FILING DATE: 1999-04-01

; EARLIER APPLICATION NUMBER: 60/080,414

; EARLIER FILING DATE: 1998-04-02									
; EARLIER APPLICATION NUMBER: 60/178,851									
; EARLIER FILING DATE: 1998-10-26									
; NUMBER OF SEQ ID NOS: 29									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 2									
; LENGTH: 2937									
; TYPE: DNA									
; ORGANISM: Pasteurella multocida									
PCT-US99-07289-2									
Query Match									
Best Local Similarity 76.6%; Score 2283.4; DB 1; Length 2937;									
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;									
Qy	43	ATTTTAAAGGAAGAAATGAATACATATATACACAGCAATAAAGCATATACAGCAAT	102						
Db	1	ATTTTAAAGGACAGAAATGAATACATATATACAGCAATAAAGCATATACAGCAAT	60						
Qy	103	GACTATGAATAGCACTCAAAATATTGAGAACTCTGCTGAACCTACGGCGGAAATC	162						
Db	61	GACTATCAATAGCACTCAAAATATTGAAAGTCGGCGGAATCTATGACGGAAAT	120						
Qy	163	GTGGAATCCAAATATCAATGTAAAGAAAACTCTGACCAATTC-----	209						
Db	121	GTGGAATTCAAATACCAATGCCAAGAAACTCTCAGCACATCCTTCTGTTAATTC	180						
Qy	210	-----TTATGTAAGTGAAGATAAATAAAGAGTGTGGATAGCTCATTTAGATATC	261						
Db	181	GCACATCTTCTGTAATAAAGAAAGAAAGTCAATGTTGGATAGTCCTGTTAGATAT	240						
Qy	262	GCACACAGCTCTTACTTCCACGTAATAAATAAATTAACCTATCCGAATCAGAAAAAC	321						
Db	241	GCACACAACTGTACTTCCACGTAATAAATAAATTAAGTACTTCTGACTCGGAAAAAC	300						
Qy	322	AGTTTAAAAAATAATGGAATCTATCACTGCGGAAAAATCGGAGACCGCAGAAATC	381						
Db	301	ACGTTAAAAAATAATGGAATGCTCACTGAGAAGAAATCTCAAAATCGGAGGTAAG	360						
Qy	382	AGGTGGAAGTACTACCAAGATTTTCTTAAGATCTGTTCTGCTCCATTCGCCAGAT	441						
Db	361	CGCGTCGCCCTTACCAAAAGATTTTCCCAAGATCTGTTTAGCGCCTTTACCTGAT	420						
Qy	442	CATGTTAATGATTTACATGGTCAAAAATCGAAAAAAGCTTAGTATAAAGCCTGTA	501						
Db	421	CATGTTAATGATTTACATGGTACAAAACGGAAGAAAGCTTGCATATAAACCTGAA	480						
Qy	502	AATAAGAAATATCGCTCTTCTATATTATTCCTACATTTAATCGTACCGCTATTTAGAT	561						
Db	481	CATCAACATATGCTCTTCTATATGTTTACAACTTCAATCGACCGCAATTTTATCG	540						
Qy	562	ATACGTTAGCCTGTTGGTCAATCAGAAACAAACTACCCTTTGAAGTCGTTGTGCA	621						
Db	541	ATTACATAGCCCTTTAGTAAACCAAAAACACATTTACCCTGTTGAAGTTATCGTGACA	600						
Qy	622	GATGATGGTAGTAAAGAAACTTACTTACCATTTGTGCAAAAATACGAACAAACTTCAC	681						
Db	601	GATGATGGTAGTCAGGAAGATCTATCACCGCATCTCGCCAATATGAAATATAATTTGGAT	660						
Qy	682	ATAAGTATGTAGACAAAAGATATTAGGATATCAATTTGTGCGATCAGAAACTTAGT	741						
Db	661	ATTGCGTACGTCAGACAAAAGATTAACGGTTTCAAGCCAGTCGGCTCGGAATATGGGA	720						
Qy	742	TTAGTCACAGCAAGATATTTTGTCTCGATCTAGACTGCGATATGGCACCAACAACAA	801						
Db	721	TTAGCCTTAGCAAAATATGACTTTATTTGGCTTACTCGACTGTGATATGGCCCAATCCA	780						
Qy	802	TTATGGGTTCACTTCTTACAGAACTATTAGAGAAATGATATGTTTAAATGGA	861						
Db	781	TTATGGGTTCACTTCTTATGTTGACAGCTATTAGAAGATGATGATTTAAACAATCATTTGGT	840						
Qy	862	CCTAGAAAATATGTCGATCTCAATATATACCGCAGACAATTCCTTAACGATCCATAT	921						

Db 841 CCAAGAAATACATCGATACAAACATATTGACCCAAAAAGACTTCTTAATAACGGAGT 900
Qy 922 TTAATAGATCACTACCTGAACCCGTACAATAACAACTCCCTCGATTACATCAAAAGGA 981
Db 901 TTGCTTGAATCATATTACAGAAGTGAAGCAATAATAGTTGGCCGCAAAAGGGGAAGA 960
Qy 982 AATATATCGTTGGATTGGAGATTAGAACATTTTCAAAAAACCCGATATCTACGCTCTATGT 1041
Db 961 ACAGTTCTCGATTGGCGCTTAGAACATTCGAAAAACAGAAATCTCCGCTTATCC 1020
Qy 1042 GATTTCCGTTTCGTTATTTTGTGGGGTAATGTTGCAATTTTCTAAGAAATGGCTAAAT 1101
Db 1021 GATTCGCTTTCCGTTTGTGGCGGGTAATGTTGCTTTTCGCTAAAAAATGGCTAAAT 1080
Qy 1102 AAGTATGTTGTTTCGATGAAGAAATTAATCATTTGGGGGGCGAAGATGTPAGAAATTTGGT 1161
Db 1081 AAATCGGGTTCTTTGATGAGGAATTAATCACTGGGGTGGGAAGATGTTGGA 1140
Qy 1162 TACAGATTATTTGCCAAAGGCTGTTTTTTCAGAGTAATTTGACGGCGGAATGGCCATCCAT 1221
Db 1141 TATCGCTTATTCGTTACGGTAGTTCTTTAAACATATTGATGGCATTAATGGCCTACCAT 1200
Qy 1222 CAAGAACCCACTGGTAAAGAAATGAACAGACGGAGCTGGTAAAGATATTACGCTT 1281
Db 1201 CAAGAGCCACCGTAAAGAAATGAACCGGATCGTGAAGCGGGAATAATATTACGCTC 1260
Qy 1282 AAAATTGTGAAGAAAGGTAACCTTACATCTATAGAAGCTTTTACCATAGAAAGATTCA 1341
Db 1261 GATATTATGAGAAAGGTCCTTATCTATAGAAAATTTTACCATAGAAAGATTGCG 1320
Qy 1342 CATATTCATAGAAATACCTTTAGTTCTTATTTATATCCCGCTTATATACTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATCCAGCTTATACTGTGCAAACTAT 1380
Qy 1402 ATTCAAAGATGTGATAGTAGTCTTAAACAACACTGTTGCGATCTCGAGGTTGTATT 1461
Db 1381 ATTCAGGTTGCGTATAGTAGTGCCTGAATCAGAGCTGTTGTATCTCGAGGTTGTATT 1440
Qy 1462 TGTAAACATGGTTCAACAGATAATACCTTAGAAGTGTATCAATAAGCTTTTATGTAATAAT 1521
Db 1441 TGTAAACATGGTTCAACAGATAATACCTTAGAAGTGTATCAATAAGCTTTTATGTAATAAT 1500
Qy 1522 CCTAGGTTACGATCATGCTTAACCAAAATGGCGGAATAGCCTCAGCATCAAAATGCAGCC 1581
Db 1501 CCTAGGTTACGATCATGCTTAACCAAAATGGCGGAATAGCCTCAGCATCAAAATGCAGCC 1560
Qy 1582 GTTTCTTTTGTAAAGTTATTACATTGGGCGATTAGATTTCAGATGATTATCTTGAGCCT 1641
Db 1561 GTTTCTTTTGTAAAGTTATTACATTGGGCGATTAGATTTCAGATGATTATCTTGAGCCT 1620
Qy 1642 GATCAGTTGAACCTGTGTTTAAAAGAAATTTTAAAGATAAAACGCTAGCTTGTTTAT 1701
Db 1621 GATCAGTTGAACCTGTGTTTAAAAGAAATTTTAAAGATAAAACGCTAGCTTGTTTAT 1680
Qy 1702 ACCACTAATAGAAAGCTCAATCCGGATGGTAGCTTAATCGCTAATGGTTACAATGGCCA 1761
Db 1681 ACCACTAATAGAAAGCTCAATCCGGATGGTAGCTTAATCGCTAATGGTTACAATGGCCA 1740
Qy 1762 GAATTTTCAACGAGAAAACTCACACGGCTATGATTGCTCACCATTATTTAGAAATGTTTACG 1821
Db 1741 GAATTTTCAACGAGAAAACTCACACGGCTATGATTGCTCACCATTATTTAGAAATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTAACGGATGGATTAAACGAAATATTGAAACGGCGTGATATTAT 1881
Db 1801 ATTAGAGCTTGGCATTTAACGGATGGATTAAACGAAATATTGAAATGGCGTAGACATAT 1860
Qy 1882 GACATGTTCCCTTAACCTCAGTGAAGTGTGAAATTTAAACATCTTAATAAAATCTGCTAT 1941
Db 1861 GACATGTTCCCTTAACCTCAGTGAAGTGTGAAATTTAAACATCTTAATAAAATCTGCTAT 1920
Qy 1942 AACCCGGTATTACATGGTGAATACATCAATTAAGAAATCGCGATTCAAAAGAAAAAC 2001
Db 1921 AACCCGGTATTACATGGTGAATACATCAATTAAGAAATCGCGATTCAAAAGAAAAAC 1980

Qy 2002 CATTTTGTGTAGTCAATCAGTCAATTAATAAGACAGGCATCAATTTATTAATTTATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATAAGACAGGCATCAATTTATTAATTTATGAC 2040
Qy 2062 AATTTGATGATTTAGATGAAGTAGAAGATATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAGTAGAAGATATATTTTCAATAAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAATGATATGTTAAAAGATCTTAAAGATCTTAAAGATCAATTAAGATGCCAAATCGCA 2181
Db 2101 GAAGATGATGATTTTAAAGATTTTAAATCATCCAGATTAAGATGCCAAATCGCA 2160
Qy 2182 GTCAGTATTTTATCCCAATACATTAACCGGCTTAGTGAAGAACTTAACAAATATTAT 2241
Db 2161 GTCAGTATTTTATCCCAATACATTAACCGGCTTAGTGAAGAACTTAACAAATATTAT 2220
Qy 2242 GAATTAATAAATAATATATTTCGTTATTATTTCTACATCTTGTATAGAATCATCTTTACACCA 2301
Db 2221 GAATTAATAAATAATATATTTCGTTATTTCCTACATCTTGTATAGAATCATCTTTACACCA 2280
Qy 2302 GACATCAAAAAAGAAATATTGGCTTTCTATCATAGCACCAAGTCAATTTTACTAAAT 2361
Db 2281 GATATCAAAAAAGAAATATTAGCTTTCTATCATAAACATCAAGTGAATTTTACTAAAT 2340
Qy 2362 AATGACATCTCATATATACAGAGTAATAGACTAATAAAGCTGAGGCACATTTAAGTAAT 2421
Db 2341 AATGATATCTCATATATACAGAGTAATAGACTAATAAAGCTGAGGCACATTTAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAAGTTAAATCTAAATGTGAATACATCAATTTTGTATAAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAAGTTAAATCTAAATGTGAATACATCAATTTTGTATAAATCATGAC 2460
Qy 2482 AGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTGGCATGAAT 2541
Db 2461 AGCCTATTCGTTAAAAATGACAGCTATGCTTATATGAAAAAATATGATGTGGCATGAAT 2520
Qy 2542 TTCACGATTTAAACACATGATTGGATCCAGAAATCAATCGCATCCACCATTTAAAAAG 2601
Db 2521 TTCACGATTTAAACACATGATTGGATCCAGAAATCAATCGCATCCACCATTTAAAAAG 2580
Qy 2602 CTGATTTAAACCTATTTTAAATGACAATGACTTAAGAAATATGAATGTGAAAGGGGCATCA 2661
Db 2581 CTGATTTAAACCTATTTTAAATGACAATGACTTAAGAAATATGAATGTGAAAGGGGCATCA 2640
Qy 2662 CAAGTATGTTTATGAAGTATGCGCTACCGATGAGCTTCTGACGATTTTAAAGAAAGTC 2721
Db 2641 CAAGTATGTTTATGAAGTATGCGCTACCGATGAGCTTCTGACGATTTTAAAGAAAGTC 2700
Qy 2722 ATCATCTCTGCAATCAATTAATGATGTGCGAGAAATATAACACTGAGGATATTTGGTTC 2781
Db 2701 ATCATCTCTGCGAGTCAATTAATGATGTGCGAGAAATATAACACTGAGGATATTTGGTTC 2760
Qy 2782 CAATTTGACCTTTTAAATCTTAGAAAAAGAAACCGGCCATGTTATTTAATAAAACATCGACC 2841
Db 2761 CAATTTGACCTTTTAAATCTTAGAAAAAGAAACCGGCCATGTTATTTAATAAAACATCGACC 2820
Qy 2842 CTGACTTATATGCTTTGGGAGCAAAATTAACATGACAAATGAACAAATTTCAAAGTGA 2901
Db 2821 CTGACTTATATGCTTTGGGAGCAAAATTAACATGACAAATGAACAAATTTGAAAGTGA 2880
Qy 2902 AAAAAAGGCAAAATATCCCGTTAAACAGTTTCAATTTAATAGTATAACGCTATAA 2958
Db 2881 AAAAGGGGAGAAATATACCTTTAAACAAGTTTCAATTTAATAGTATAACTCTATAA 2937

RESULT 8

US-09-469-200-2

; Sequence 2, Application US/09469200

; GENERAL INFORMATION:

; APPLICANT: Board of Regents of the University of Oklahoma

; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS

; TITLE OF INVENTION: OF USE

FILE REFERENCE: 617022-7
CURRENT APPLICATION NUMBER: US/09/469,200
CURRENT FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/080,414
PRIOR FILING DATE: 1998-04-02
PRIOR APPLICATION NUMBER: 60/178,851
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 2937
TYPE: DNA
ORGANISM: Pasteurella multocida
US-09-469-200-2

Query Match 76.68; Score 2283.4; DB 18; Length 2937;
Best Local Similarity 86.58; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;

Qy 43 ATTTTAAAGGAAAGAAATGAATACATATTACAAAGCAATATAAGCATATAACAGCAAT 102
Db 1 ATTTTAAAGGACAGAAATGAATACATATTACAAAGCAATATAAGCATATAACAGCAAT 60

Qy 103 GACTATGAATAGCACTCAAAATTTATTTGAGAAGTCTGCTGAACCTACGGCGGAAAAATC 162
Db 61 GACTATCAATTAGCACTCAAAATTTATTTGAAAAGTCGGCGGAAATCTATGGACGGAAAAAT 120

Qy 163 GTTCAATTCCAAATTTACAAATGTAAAGAAAAAATCTCGACCAATTC----- 209
Db 121 GTTGAATTTCAAATTTACCAATGCCAAGAAAAAATCTCAGCACATCTCTTGTGTTAAATCA 180

Qy 210 -----TTATGTAAGTGAAGATAAAAAAACAAGTGTTCGATAGCTCATATAGATATC 261
Db 181 GCACATCTTCTGTAATAAAGAGAAAAAGTCAATGTTTCGGATAGTCGCTAGATATT 240

Qy 262 GCAACAGCTCTTACTTTCACAGTAAAAAATAAATTAACCTATCCGGAATCAAAAAAAC 321
Db 241 GCAACACAACCTGTACTTTCACAGTAAAAAATAAATTAAGTACTTCTGACTCGGAAAAAAC 300

Qy 322 AGTTTAAAAAATAATGAAATCTATCACTCGGAAAAAATCGAGAACGCAAAATCAGA 381
Db 301 ACGTTAAAAAATAATGAAATCTGCTACTGAGAGAAATCTGAAATCGCGAGGTAAGA 360

Qy 382 AAGGTGAACCTAGTACCCAAAGATTTTCCTAAAGATCTTGTTCCTGCCATTCGCCAGAT 441
Db 361 GCGGTGCGCCTTGTACCAAAAGATTTCCCAAGATCTGTTTTCGCGCCTTTACCTCAT 420

Qy 442 CATGTTAATGATTTTACATGTTACAAAATCGAAAAAAGCTTAGGTATAAAGCGCTGTA 501
Db 421 CATGTTAATGATTTTACATGTTACAAAAGCAAAAGCAAAAGACTTGGCATAAAAACCTCAA 480

Qy 502 AATAAGATATCGCTCTTCTATTATTATTTCTACATTAATCTAGCTAGCCGTATTAGAT 561
Db 481 CATCAACATGTGGTCTTCTTATTATCGTTTACAACTCAATCGACCACTTTTATCG 540

Qy 562 ATACGTTAGCTGTTGGTCAATCAGAAAAACAACCTACCCATTTGAAAGTCTGTTGCA 621
Db 541 ATTACATTAGCTGTTTAGTAACCAAAAACACATTACCCGTTTGAAGTTATCGTGACA 600

Qy 622 GATGATGTAGTAAGAAAACTTACTTACCANTGTGCAAAATAGCAAAAAAATCTTGAC 681
Db 601 GATGATGTAGTACAGGAAGATCTATCACCGATCATCTGCCAATATGAAAAATAAATGGAT 660

Qy 682 ATAAAGATGTAAACAAAAAGATTATGGATATCAATTTGTGTCAGTCAAGAACTTAGT 741
Db 661 ATTCGCTACGTCAGCAAAAAAGATAACGGTTTTCAAAGCCAGTGGCGCTCGGAATATGGGA 720

Qy 742 TTACGCTACGCAAAAGATGATTTTGTCTCGATTTCTAGACTGCGATATGGCAACCAACAA 801
Db 721 TTACGCTTAGCAAAATATGACTTTATTTGGCTTACTCGACTGTGATATGGCGCAATCCA 780

Qy 802 TTATGGGTTCATCTTATCTTACAGAACTATTAGAAGACAATGATATTGTTTAAATGGA 861
Db 1861 GACATGTTCTCTCAAACTCAGTGAAGTTGAAAAATTTAAACATCTTAAATAAATCTGCTAT 1920

Db 781 TTATGGGTTCATTTCTTATGTTTCAGAGCTATTAGAAGATGATGATTTAAACAATCATTTGCT 840

Qy 862 CCTAGAAAATATGTCGATCTCATATATTACCGAGAGCAATTTCTTAAACGATCCATAT 921
Db 841 CCAAGAAAATATCATGATACACAACATATTGACCCAAAAGACTTCTTAAATACCGGAGT 900

Qy 922 TTAATAGAATCACTACCTGAAACCGCTACAAATAACAATACCTTCGATTTACATCAAAAGGA 981
Db 901 TTGCTTGAATCATTACAGAGAGTGAACCAATATAAGTGTGGCCGCAAAAGGGAAGA 960

Qy 982 AATATATCGTTGGATTTGGAGATTAGAACAATTTCAAAAAACCGAATATCTAGCTCTATGT 1041
Db 961 ACAGTTTCTCTGGATTGGCGCTTAGAACAAATTCGAAAAAACAGAAAATCTCGCTTATCC 1020

Qy 1042 GATTCCTCGTTTCGTTATTGTTGTCGGGTAATGTTGCAATTTCTAAAGAAATGGCTAAAT 1101
Db 1021 GATTCGCCCTTTCCGTTTTCGCGGGGTAATGTTGCTTTCGCTAAAAAATGGCTAAAT 1080

Qy 1102 AAAAGTAGGTTCGATGAAGAAATTTAATCATTTGGGGGGGGAAGATGATAGAATTTGGT 1161
Db 1081 AAATCCGGTTTCTTGTATGAGGAATTTAATCACTGGGTTGGAGAGATGTGGAATTTGGA 1140

Qy 1162 TACAGATTTATTTGCCAAAGGCTGTTTTCAGAGTAATTTGACGGGGAATGGCCATCCAT 1221
Db 1141 TATCGCTTATTCGTTACGGTAGTTTCTTTAAAACTATTGATGGCAATTATGGCTACCAT 1200

Qy 1222 CAAGAACCACTGCTAAAGAAAATGAACAGAACGCGAAGCTGGTAAAAAGTATTACGCTT 1281
Db 1201 CAAGAACCACTGCTAAAGAAAATGAACCGGATCGTGAAGCGGGAAGAAAATATTACGCYC 1260

Qy 1282 AAAATTTGTAAGAAAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAAGATTCA 1341
Db 1261 GATATTATGAGAGAAAAGGTCCCTTATATCTATAGAAAATTTTACCAATAGAAGATTCTG 1320

Qy 1342 CATATTTCATAGAAATACCTTTAGTTTCTATTATATCCCGCTTATAAATCTGTCGAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATCCCACTTATAAATCTGTCGAAATAT 1380

Qy 1402 ATTCAAGATGTAGATGCTCTTAAATCAAACTGTTGTCGATCTCAGAGTTTGTATT 1461
Db 1381 ATTCAAGTGTGGTAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440

Qy 1462 TGTAAAGATGTTCACAGATAATACCTTTAGAAGTGAATCAATGAAGCTTTTATGGTAATAAT 1521
Db 1441 TGTAAAGATGTTCACAGATAATACCTTTAGAAGTGAATCAATGAAGCTTTTATGGTAATAAT 1500

Qy 1522 CCTAGGTTACGATCATGCTTAAACCAATGGCGGAATAGCTTCAGCATCAATAGCAGCC 1581
Db 1501 CCTAGGTTACGATCATGCTTAAACCAATGGCGGAATAGCTTCAGCATCAATAGCAGCC 1560

Qy 1582 GTTCTCTTTGCTAAAGGTTATTACATTTGGGAGTATGATTCAGATGATTTATCTTGAGCCT 1641
Db 1561 GTTCTCTTTGCTAAAGGTTATTACATTTGGGAGTATGATTTAGATGATTTATCTTGAGCCT 1620

Qy 1642 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAAACGCTAGCTGTGTTTAT 1701
Db 1621 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAAACGCTAGCTGTGTTTAT 1680

Qy 1702 ACCACTAATAGAAAGCTCAATCCGATGCTAGCTTAATCGTAAATGGTTACAAATGGCCA 1761
Db 1681 ACCACTAATAGAAAGCTCAATCCGATGCTAGCTTAATCGTAAATGGTTACAAATGGCCA 1740

Qy 1762 GAATTTTTCAGGAAAAAATCAACGGCTATGATGCTCACCATTGTTAGAAATGTTTACG 1821
Db 1741 GAATTTTTCAGGAAAAAATCAACGGCTATGATTTGCTCACCCTTTTAAATGTTTACG 1800

Qy 1822 ATTAGAGCTTGGCAATTTAAACGGATTTAAACGAAAAATTTGAAACCCGCTGGATAT 1881
Db 1801 ATTAGAGCTTGGCAATTTAACTGATGATCAATGAAAAATTTGAAAAATCCGCTAGACTAT 1860

Qy 1882 GACATGTTCTTAACTCAGTGAAGTTGAAAAATTTAAACATCTTAAATAAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTCAGTGAAGTTGAAAAATTTAAACATCTTAAATAAATCTGCTAT 1920

QY 1942 AACCCGGTATTACATGGTGATTAACACATCCATTAGAAACTCGGCATTCAAAAGAAAAC 2001
Db 1921 AACCCGTATTATCATGGTGATTAACACATCAATTAAGAACTTTGGCATTCANAAGAAAAC 1980
QY 2002 CATTTTGTGTAGTCAATCAGTCATTAAATACAGAGGCATCAATTTATTATAATTATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCATTAAATACAGAGGCATCAATTTATTATAATTATGAC 2040
QY 2062 AAATTTGATGATTTAGATGAAAGTAGAAGTATATCTTCAATTAACACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAAGTAGAAGTATATCTTCAATTAACACCGCTGAATATCAA 2100
QY 2122 GAAGAAATCGATGATTTAAAGATCTTAAACTCAATTCARAAATAAAGATGCCAAATCGCA 2181
Db 2101 GAAGAGATTTGATATCTTAAAGATATTAATAATCATCCAGAAATAAAGATGCCAAATCGCA 2160
QY 2182 GTCAGTATTTCTATCCCAATACATTAACCGCTTAGTGAAGAACTAAACAAATATTAAT 2241
Db 2161 GTCAGTATTTCTATCCCAATACATTAACCGCTTAGTGAAGAACTAAACAAATATTAAT 2220
QY 2242 GAATATAATAAATAATATTCGTTATTAATCTACATGTTGATGAAGATCATCTTACACCA 2301
Db 2221 GAATATAATAAATAATATTCGTTATTAATCTACATGTTGATGAAGATCATCTTACACCA 2280
QY 2302 GACATCAAAAAGAAATATTGGCTTTCTATCATAAGCACCAAGTGAATATTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTGGCTTTCTATCATAAGCACCAAGTGAATATTTTACTAAAT 2340
QY 2362 AATGACATCTCATATTACAGGATATAGACTAATAAATACTGAGGCACATTTTAAAGTAAT 2421
Db 2341 AATGATATCTCATATTACAGGATATAGATTAAATAAATACTGAGGCACATTTTAAAGTAAT 2400
QY 2422 ATTAATAAATAAGTACATTAATCTAATTTGATGATACATCAATTTTGTATAATCATGAC 2481
Db 2401 ATTAATAAATAAGTACATTAATCTAATTTGATGATACATCAATTTTGTATAATCATGAC 2460
QY 2482 AGCTATTCGTTAAAATGACAGTATGCTTATATGAAAAATATGATGCGGCATGAAT 2541
Db 2461 AGCTATTCGTTAAAATGACAGTATGCTTATATGAAAAATATGATGCGGCATGAAT 2520
QY 2542 TTCTCAGCATTAACACATGATTTGATGAGAGAAATCAATCGGCATCCACCATTTAAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTTGATGAGAGAAATCAATCGGCATCCACCATTTAAAAG 2580
QY 2602 CTGATTTAAACCTATTTTATGACATGACTTAAGAGATGATGATGAAAGGGCATCA 2661
Db 2581 CTGATTTAAACCTATTTTATGACATGACTTAAGAGATGATGATGAAAGGGCATCA 2640
QY 2662 CAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACGATTTAATAAGAAAGTC 2721
Db 2641 CAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACGATTTAATAAGAAAGTC 2700
QY 2722 ATCACATCTGCCATCAATTTATGATGATGTCGAGAAATATACACTGAGGATATTTGGTTC 2781
Db 2701 ATCACATCTTTGCCATCAATTTATGATGATGTCGAGAAATATACACTGAGGATATTTGGTTC 2760
QY 2782 CAATTTGCACCTTTTAACTTTAGAAAAGAAAACCGCCATGATTTAATAAACAATCGACC 2841
Db 2761 CAATTTGCACCTTTTAACTTTAGAAAAGAAAACCGCCATGATTTAATAAACAATCGACC 2820
QY 2842 CTGACTTTATATGCTTTGGGACGAAAATTAACAATGGACAAATGAACAAATTTCAAAGTGCA 2901
Db 2821 CTGACTTTATATGCTTTGGGACGAAAATTTACAATGGACAAATGAACAAATTTGAAAGTGCA 2880
QY 2902 AAAAAAGCGGAAAATATCCCGTTTAAACAAGTTCAATTTATATGATTAACCTATAA 2958
Db 2881 AAAAAAGCGGAAAATATACCTGTTTAAACAAGTTCAATTTATATGATTAACCTATAA 2937

; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUITIMILIS HYALURONAN SYNTHASE GENE AND EXPRESS
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: pasteurella multocida
US-09-879-959-9

Query Match 76.6%; Score 2283.4; DB 33; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;
QY 43 ATTTTAAAGGAAGAAATGAATACATTATCACAGCAATAAAGCATATAACGCAAT 102
Db 1 ATTTTAAAGGACAGAAATGAATACATTATCACAGCAATAAAGCATATAACGCAAT 60
QY 103 GACTATGAATTAGCAGCTCAATTTATTTGAGAGCTCTGCTGAAACCTACGGCGGAAATC 162
Db 61 GACTATCAATTAGCAGCTCAATTTATTTGAAAGTCGGCGGAAATCTATGGACGGAAT 120
QY 163 GTTGAATTCACAAATTTACAAATGTAAAGAAAAAATCTCGACCAATTC----- 209
Db 121 GTTGAATTTCAATTTACCAATGCCAAGAAAACTCTCAGCACATCTCTCTGTTAATTC 180
QY 210 -----TTATGTAAGTGAAGATAAATAAACAAGTGTTCGGATAGTCTAATAGATC 261
Db 181 GCACATCTTCTGTAAATAAAGAAAGAAAGTCAATGTTTCGATAGTTCGTTAGATAT 240
QY 262 GCACACAGCTCTTACCTTCCACAGTAAATAAATAATTAATCTCTATCGAATCAGAAAAAAC 321
Db 241 GCAACACACTGTTACTTTCACAGCTAAATAAATAATAGTACTTCTGACTCGGAAAAAAC 300
QY 322 AGTTTAAAAATAAATGAAATCTATCACTGGGAAAAAATCGGAGAACGAGAAATCAGA 381
Db 301 ACGTTAAAAATAAATGAAATTCCTCAGTGAAGAAATCTGAAAAATCGGAGGTAAGA 360
QY 382 AAGTGAAGTACTACCCAAAGATTTTCCTAAAGATCTTGTCTTGCTCCATTTGCCAGAT 441
Db 361 GCGTCCGCCCTGTACCAAAAGATTTTCCCAAGATCTGGTTTAGCCCTTTACCTGAT 420
QY 442 CATGTTAATGATTTTACATGTCACAAAAATCGAAAAAAGCTTAGTATATAAGCCTGTA 501
Db 421 CATGTTAATGATTTTACATGTCACAAAAAGGAAAAAGACTTGGCATAAAACCTGAA 480
QY 502 AATAAGAAATCGGCTCTTCTTATTTATTTATTCCTACATTTAATCGTAGCGGTATTTAG 561
Db 481 CATCAACATGTTGGTCTTCTTATTTATTCGTTACAACTTCGACCAAGCAATTTTATCG 540
QY 562 ATAAGTGTAGCTGTTTGTGTCATCAATCAGAAAAAACAATACCCTATTGAGTCTGTTGCA 621
Db 541 ATTACATTAGCTGTTTGTAGTAAACCAAAAAACACATTACCCCTTTGAGTTCGTGACA 600
QY 622 GATGTTGTAGTGAAGAAACTTACTTACCATTGTGCAAAAAATACGACAAAACTTGAC 681
Db 601 GATGATGTAGTCAAGGAAGATCTATCACCGATCATTCGCAATATGAAAAATAAATGGAT 660
QY 682 ATAAAGTATGAAGACAAAAAGATTTATGATATCAATTTGTGTGAGTCAAGAACTTAGT 741
Db 661 ATTGCTACGTACAGACAAAAAGATAACGGTTTTCAGCGCAGTGGCGCTCGGAATATGGGA 720

Qy 742 TTACGTACAGCAAGATGATGATTTTGTCTCGATCTTCTAGATCGGATATGCGACCAACA 801
Db 721 TTACGCTTAGCAAAATATGACTTTTATTTGCTTACTCGATCTGATATGCGCCAAATCCA 780
Qy 802 TTATGGGTTTCATCTTATCTTACAGCAATCTTAGAAGACATGATATGTTTAAATTTGA 861
Db 781 TTATGGGTTTCATCTTATGTTGACAGCTATTTAGAAGATGATGATTTAAACATCATTTGGT 840
Qy 862 CTTAGAAAATTTGGATGATCAATATATACCGCAGAACATTTCTTAAACGATCCATAT 921
Db 841 CCAAGAAATACATCGATACACAATATTTGACCCAAAGACTCTTAAATTAACGGGAT 900
Qy 922 TTAATPAGAACTACCTGAAACCGCTACAAATACAAATCTCTGATACATCAAAAGGA 981
Db 901 TTGCTTGAATCATTTACAGAAAGTGAACCAATATAGTTGTCGCAAAAGGGGAAGGA 960
Qy 982 AATATATCGTTGGATGAGATAGAACATTTCAAAAACCGATTAATCTAGCTATGCT 1041
Db 961 ACAGTTTCTCGATTTGGCGCTTAGAACATTTGGAACAAATTCGAAACAAAGCAATCTCCGCTTATCC 1020
Qy 1042 GATTCCTCGTTTCGTTATTTTGTTCGGGTAAATGTTGATTTTCTTAAAGATGGCTAAAT 1101
Db 1021 GATTCGCTTTCCGTTTTCGCGGGGTAAATGTTGCTTTCGCTTAAAAATGGCTAAAT 1080
Qy 1102 AAAGTAGGTTGGTTCGATCAAGAAATTTAATCATTTGGGGGGGGAAGATGTAGAAATTTGGT 1161
Db 1081 AAATCCGGTTCTTTGATGAGGAATTTAATCATCTGGGTGGAGAGATGTGGAATTTGGA 1140
Qy 1162 TACAGATATTTGCAAAAGGCTGTTTTTTCAGAGTAAATGACGGGGGAATGCCATTCAT 1221
Db 1141 TATCGCTTATTTCCGTTACGGTAGTTCTTTTAAACATATTTGATGGCATTTATGCCATACCAT 1200
Qy 1222 CAAGAACCACCTGGTAAAGAAAATGAAACAGAACCGGAAGCTGGTAAAGATTTAGCGTT 1281
Db 1201 CAAGAGCCACCGAGTAAAGAAAATGAAACCGGATCTGGAAGCGGGGAAATATTTACGCTC 1260
Qy 1282 AAAATTTGGAAGAAAAGTACCTTACATCTATAGAAAGCTTTTACCATAAGAAATCCA 1341
Db 1261 GATATTATGAGAAAAGTCCCTTATATCTATAGAAAATTTTACCATAAGAAATTCG 1320
Qy 1342 CATATTCATAGAAATACCTTTAGTTTCTATTTATATCCCGCTTTTAACTGTGCAATTTAT 1401
Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATCCCGAGCTTATAAATGTGCAAACTAT 1380
Qy 1402 ATTCAGAGTGTAGATAGTCTCTTAAATCAAACTGTTGTCGATCTCGAGTTTGTATTT 1461
Db 1381 ATTCAGCTGCGTAGATAGTGCATGATCAGATGTTGTTGTTGCTCGAGGTTTGTATTT 1440
Qy 1462 TGTAAAGATGTTCAACAGATAATACCTTAGAAGTGCATCAATAGGCTTTATGGTAAATAT 1521
Db 1441 TGTAAAGATGTTCAACAGATAATACCTTAGAAGTGCATCAATAGGCTTTATGGTAAATAT 1500
Qy 1522 CTTAGGTTACGATCATGTCTAAACCAATGCGGGAATAGCCCTCAGCATCAATGCGAGCC 1581
Db 1501 CTTAGGTTACGATCATGTCTAAACCAATGCGGGAATAGCCCTCAGCATCAATGCGAGCC 1560
Qy 1582 GTTCTCTTTTGTCTTAAAGGTTATTTACATTTGGCAGTTAGATTTAGATTTAGATTTAGATTT 1641
Db 1561 GTTCTCTTTTGTCTTAAAGGTTATTTACATTTGGCAGTTAGATTTAGATTTAGATTTAGATTT 1620
Qy 1642 GATCAGTTGAACTGTGTTTAAAGAAATTTTAAAGATTAACACGCTAGCTTTGTGTTTAT 1701
Db 1621 GATCAGTTGAACTGTGTTTAAAGAAATTTTAAAGATTAACACGCTAGCTTTGTGTTTAT 1680
Qy 1702 ACCACTAATAGAAAGCTCAATCCGGATGTTAGCTTAAATGCGTAAATGTTTACAAATTTGCGCA 1761
Db 1681 ACCACTAATAGAAAGCTCAATCCGGATGTTAGCTTAAATGCGTAAATGTTTACAAATTTGCGCA 1740
Qy 1762 GAATTTTTCAGGAGAAAACCTCACACGCTATGATTTGCTCACCATTTTAGAATGTTTAGC 1821
Db 1741 GAATTTTTCAGGAGAAAACCTCACACGCTATGATTTGCTCACCATTTTAGAATGTTTAGC 1800
Qy 1822 ATTAGAGCTTGGCATTTTAAACGGATGGATTTTAAACGAAATATTTGAAACCGCTGGATTTAT 1881

Db 1801 ATTAGAGCTTGGCATTTAACTGATGATCAATGAAATAATTCAAAATGCGGTAGACTAT 1860
Qy 1882 GACATGTTCCCTTAAACTCAGTGAAGTTGAAATAATTTAAACATCTTTAAATAATCTGCTAT 1941
Db 1861 GACATGTTCCCTTAAACTCAGTGAAGTTGAAATAATTTAAACATCTTTAAATAATCTGCTAT 1920
Qy 1942 AACCGGTTATTTACATGCTGATTAACACATCCATTAAGAAACTCGCATTTCAAAAGAAAAC 2001
Db 1921 AACCGGTTATTTACATGCTGATTAACACATCAATTAAGAAACTCGCATTTCAAAAGAAAAC 1980
Qy 2002 CATTTTGTGTAGTCAATCAGTCAATTAATPAGACAAAGGATCAATTTATTAATTAATGAC 2061
Db 1981 CATTTTGTGTAGTCAATCAGTCAATTAATPAGACAAAGGATCAATTTATTAATTAATGAC 2040
Qy 2062 AAATTTGATGATTTAGATGAAAGTACAAAGTATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAAGTACAAAGTATATTTTCAATAAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATGATTTAAAGATCTTAAACTCAATTAAGAAACTGCAAAATCGCA 2181
Db 2101 GAAGAGATTTGATATCTTAAAGATATTAATAATCATCCAGATTAAGATGCAAAATCGCA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTGAATAAACTTAAACAATATTTAT 2241
Db 2161 GTCAGTATTTTCTATCCCAATACATTTAAACGGCTTAGTGAATAAACTTAAACAATATTTAT 2220
Qy 2242 GAATATAATAAAATATATTCGTTATTTTCTACATGTTGATTAAGATCAATCTTTACACCA 2301
Db 2221 GAATATAATAAAATATATTCGTTATTTTCTACATGTTGATTAAGATCAATCTTTACACCA 2280
Qy 2302 GACATCAAAAGAAATATTTGGCTTCTATCATAAAGCACCAGCTGAATATTTTACTAAAT 2361
Db 2281 GATATCAAAAGAAATATTTAGCTTCTATCATAAACATCAAGTGAATATTTTACTAAAT 2340
Qy 2362 AATGACATCTCATATTTACAGGATTAATAGACTAATAAAACTGAGGCACATTTTAAAGTAAT 2421
Db 2341 AATGATATCTCATATTTACAGGATTAATAGACTAATAAAACTGAGGCACATTTTAAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTAAATCTAATTTGTAATACATCAATTTTGTATTAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTAAATCTAATTTGTAATACATCAATTTTGTATTAATCATGAC 2460
Qy 2482 AGCTATTTCTGTTAAATAACAGCTATGCTTATATGAAATAATATGATGTCGGCATCAAT 2541
Db 2461 AGCTATTTCTGTTAAATAACAGCTATGCTTATATGAAATAATATGATGTCGGCATCAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTTGGATCGAGAAAATCAATGCGCATCCACCATTTTAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTTGGATCGAGAAAATCAATGCGCATCCACCATTTTAAAG 2580
Qy 2602 CTGATTTAAACCTTATTTTAAAGCAATGACTTAAAGATGATGAATGTAAGAGGCAATCA 2661
Db 2581 CTGATTTAAACCTTATTTTAAAGCAATGACTTAAAGATGATGAATGTAAGAGGCAATCA 2640
Qy 2662 CAAGGTATGTTTAAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAGATC 2721
Db 2641 CAAGGTATGTTTAAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAGATC 2700
Qy 2722 ATCAGATCTCGCAATCAATTTGATGATGTCGCAAGATATTAACACTGAGGATTTTGGTTC 2781
Db 2701 ATCAGATCTCGCAATCAATTTGATGATGTCGCAAGATATTAACACTGAGGATTTTGGTTC 2760
Qy 2782 CAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCATCTATTTTAAATAACATCGACC 2841
Db 2761 CAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCATCTATTTTAAATAACATCGACC 2820
Qy 2842 CTGACTTATATGCTTTGGGAAACGAAATTAATGGAATAATGAACAAATTTCAAGTGCA 2901
Db 2821 CTGACTTATATGCTTTGGGAAACGAAATTAATGGAATAATGAACAAATTTCAAGTGCA 2880
Qy 2902 AAAAAAGGGGAAATATCCCGGTTAAACAGTTTCAATTTATTAATAGTATAACGCTATAA 2958
Db 2880 AAAAAAGGGGAAATATCCCGGTTAAACAGTTTCAATTTATTAATAGTATAACGCTATAA 2958

Db 2881 AAAAGAGGAGAAATATACCTGTTAAACAAGTTCATTATTAAATAGTATAACTATATA 2937

RESULT 10
US-10-172-527-9
; Sequence 9, Application US/10172527
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS SUBT
; FILE REFERENCE: 3554.048
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US/10/172,527
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/297,788
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/297,744
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: pasteurella multocida
US-10-172-527-9

Query Match 76.6%; Score 2283.4; DB 41; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2540; Conservative 0; Mismatches 376; Indels 21; Gaps 1;

Qy 43 ATTTTAAAGGAAGAAATGAATACATTATACAAAGCAATATAAGCATATACAGCAAT 102
Db 1 ATTTTAAAGGACAGAAATGAATACATTATACAAAGCAATATAAGCATATACAGCAAT 60

Qy 103 GACTATGAATAGCACACAAATATTGAGAAAGTCTGTGAAAGCTAGCGGCGAAATAC 162
Db 61 GACTATCAATAGCACACAAATATTGAAAGTCTGTGAAAGCTATGAGCGGAAAT 120

Qy 163 GTTGAAATCCAAATATCAAAATGAAGAAACCTCTCGACCAATTC----- 209
Db 121 GTTGAAATTCAAATATCCAAATGCCAAGAAACCTCTCGACCAATTC----- 180

Qy 210 -----TTATGTAAGTGAAGATAAAGAAACACAGTGTTCGGATAGCTCATTTAGATATC 261
Db 181 GCACATCTTCTGTAAATATAAGAAAGAAAGTCAATGTTTCGGATAGTCCGTTAGATAT 240

Qy 262 GCAACACAGCTCTTACTTTCCAAAGTAAAGAAATTAACCTCTATCCGAATCAGAAAGAAAC 321
Db 241 GCAACACAGCTTACTTTCCAAAGTAAAGAAATTAAGTACTTTCTGACTCGGAAAGAAAC 300

Qy 322 AGTTTAAAGAAATGAATGAAATCTATCACTGGGAAAGAAATCGGAGAACGCGAATCAGA 381
Db 301 ACGTTAAAGAAATGAATGAAATCTGCTACTGAGAGAAATCTGAAATCGGAGGTAAGA 360

Qy 382 AAGTGAAGTACAGTACCAAGATTTTCTTAAAGATCTGTTCTGCTCCATTGCGAGAT 441
Db 361 GCGGTGCGCCCTGTATACCAAGATTTTCCCAAGATCTGTTTTCAGCCCTTACCTGAT 420

Qy 442 CATGTTAATGATTTTACATGTTTACAAAGTAAAGAAATCGAAAGAAAGCTTGGCATAAACCTGAA 501
Db 421 CATGTTAATGATTTTACATGTTTACAAAGTAAAGAAAGCTTGGCATAAACCTGAA 480

Qy 502 AATAGAAATATCGGCTCTTCTATTTATTTCTACATTTAATCTAGCTAGCGTATTTAGAT 561
Db 481 CATCAACATGTTGGTCTTCTATTTATTTCTGTTAATCAACATTTCAATCGACCGCAATTTATCG 540

Qy 562 ATAAGCTAGGCTGTTTGGTCAATCAGAAAGAAAGTACCCCAATTTGAAGTCTGTTTGA 621
Db 541 ATTACATAGGCTGTTTAGTAAACCAACAAACACATATACCCGTTTGAAGTATCGTGACA 600

Qy 622 GATGATGGTAGTAGGAAACCTTACTTACCAATTTGTCAGAAATATACGAACAAACCTTGAC 681
Db 601 GATGATGGTAGTCAGGAAGATCTATCACCAGTATCATGCCCAATATGAAATATGAT 660

Qy 682 ATAAAGTATGTAAAGCAAAAGATTTAGGATATCAATTTGTGCAGTCAGAAACCTTAGT 741
Db 661 ATTCGCTACGTCAGCAAAAGATTAACGGTTTTCAGCCAGTCGCGCTCGGAATATGGA 720

Qy 742 TTACGTACAGCAAAAGTATGATTTGTCGATTTAGACTTCGATATGCGATGCGCAACAA 801
Db 721 TTACGCTTACAAATATGACTTTTGGCTTACTCAGCTGTGATATGCGCCCAATATGGA 780

Qy 802 TTATGGGTTTCATTTCTTATCTTACAGAACTATTTAGAGCAATGATATTTTATTTGGA 861
Db 781 TTATGGGTTTCATTTCTTATGTCAGAGCTATTAGAAGATGATGATTAAACAATCATGGT 840

Qy 862 CCTAGAAAATATGTTGATCTCATATAATTTACCCAGCAAAATTTCTTAAAGATCCATAT 921
Db 841 CCAAGAAAATATGATGATACACAACTATTTGACCCAAAGACTTCTTAAATATACCGGAT 900

Qy 922 TTAATAGAATCACTTACCTGAAACCGCTACAAATTAACATCTCTCGATTTACATCAAAAGA 981
Db 901 TTGCTTGAATCATTTACCAGAGTGAACCAATTAATAGTGTTCGCGCAAAAGGGAAGA 960

Qy 982 AATATATCGTTGATGAGATTTAGAACATTTCAAAAACCGGATATCTAGCTATATGT 1041
Db 961 ACAGTTTCTCTGGATTTGGCGCTTTAGAACATTTCAAAAACCGGATATCTCGCTTATCC 1020

Qy 1042 GATTCTCGGTTTCGTTATTTGTTGCGGTAATGTTGTCATTTTCTTAAAGATGGCTAAAT 1101
Db 1021 GATTCTCGGTTTCGTTATTTGCGGCGGTAATGTTGTCGCTTAAAGATGGCTAAAT 1080

Qy 1102 AAAGTAGTGTGTTTCGATGAAGAATTTAATCATTTGGGGGGGCGAAGATGTAGATTTGGT 1161
Db 1081 AAATCGGTTTCTTTGATGAGGAATTTAATCACTGGGTGGAGAGATGTGGAATTTGGA 1140

Qy 1162 TACAGATTTATTTGCCAAAGCTGTTTTCAGAGTAATTTGAGCGGGAATGGCCATCAT 1221
Db 1141 TATCGCTTATTTCCGTTACGGTAGTTTCTTAAACTATTTGATGCGATTTATGCGCTACCAT 1200

Qy 1222 CAAGAACCCCTTGGTAAAGAAATGAAACAGAACCGAGCTGGTAAAGTATTTACGCTT 1281
Db 1201 CAAGAGCCACCCAGGTAAAGAAATGAAACCGATCTGGAACGGGAAAGAAATATACGCTC 1260

Qy 1282 AAAATGTGAAGAAAGGTACCTTACATCTATAGAAAGCTTTTACCAATAGAGATTC 1341
Db 1261 GATATTATGAGAAAGGTCCCTTATATCTATAGAAAGCTTTTACCAATAGAGATTCG 1320

Qy 1342 CATATTATAGAAATACCTTTAGTTTCTATTTATATCCCGCTTATAACTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTAGTTTCAATTTATATCCAGCTTATATCTGCAACATAT 1380

Qy 1402 ATTCAAGATGTAGATAGTGTCTTATCAAACTGTTGTCTGATCTCGAGGTTTGTAT 1461
Db 1381 ATTCACGTTGGTAGATAGTACCTGATCAGACTGTTGTTGATCTCGAGGTTTGTAT 1440

Qy 1462 TGTACAGTGTGTACAGATTAATACCTTAGAAGTATCAATAAGCTTTTATGTAATAT 1521
Db 1441 TGTACAGTGTGTTCACAGATTAATACCTTTAGAGTATCAATAAGCTTTTATGTAATAT 1500

Qy 1522 CCTAGGTACGATCATGTCTAAACCAATGCGGAATAGCCCTCAGCATCAAAATGACGC 1581
Db 1501 CCTAGGTACGATCATGTCTAAACCAATGCGGAATAGCCCTCAGCATCAAAATGACGC 1560

Qy 1582 GTTCTCTTTTGTAAAGGTTTATACATTTGGGCGAGTTAGATTCAGATGATTTCTTGAGCCT 1641
Db 1561 GTTCTCTTTTGTAAAGGTTTATACATTTGGGCGAGTTAGATTCAGATGATTTCTTGAGCCT 1620

Qy 1642 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAGCTAGCTGCTGTTTAT 1701
Db 1621 GATGAGTTGAACTGTGTTTAAAGAAATTTTAAAGATAAAGCTAGCTGCTGTTTAT 1680

Qy 1702 ACCACTAATAGAAACGTCACATCCGAGTGGTACGCTTAATCGCTTAATGTTACAAATGGCCA 1761
Db 1681 ACCACTAATAGAAACGTCACATCCGAGTGGTACGCTTAATCGCTTAATGTTACAAATGGCCA 1740
Qy 1762 GAATTTTCACGAGAAAACTCACAACGGCTATGATGCTACCAATTTAGAGATGTTTACG 1821
Db 1741 GAATTTTCACGAGAAAACTCACAACGGCTATGATGCTACCAATTTAGAGATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTAAACGGATGGATTTAACGAAAAATATTGAAACGGCGGTGATAT 1881
Db 1801 ATTAGAGCTTGGCATTTAAACGGATGGATTTAAACGGATGGATTTAAACGGCGGTGATAT 1860
Qy 1882 GACATGTTCCCTTAAACTCAGTGAAGTTGGAAAAATTTAAACATCTTAATAAAATCTGCTAT 1941
Db 1861 GACATGTTCCCTTAAACTCAGTGAAGTTGGAAAAATTTAAACATCTTAATAAAATCTGCTAT 1920
Qy 1942 AACCGGTATTTACATGTTGATGATAACACATCCATTAAGAACTCGGCATTCAAAAGAAAAAC 2001
Db 1921 AACCGGTATTTACATGTTGATGATAACACATCCATTAAGAACTCGGCATTCAAAAGAAAAAC 1980
Qy 2002 CATTTTGTGTGAGTCAATCAGTCAATTAATAGACAAGGCATCAATTTATTAATTTATGAC 2061
Db 1981 CATTTTGTGTGAGTCAATCAGTCAATTAATAGACAAGGCATCAATTTATTAATTTATGAC 2040
Qy 2062 AAATTTGATGATTTAGATGAAAGTAGAAAGTATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAAGTAGAAAGTATATTTTCAATAAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATGTTTAAAGATCTTAAAGTCAATTAAGATGATGCAAAATCGCA 2181
Db 2101 GAAGAGATTTGATATCTTAAAGATATTTAAATATCATCCAGATAAAGATGCCAAATCGCA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACAAATATTAT 2241
Db 2161 GTCAGTATTTTCTATCCCAATACATTAACGGCTTAGTGAAGAACTAAACAAATATTAT 2220
Qy 2242 GAATATAATAAAATATATTCGTTATTATCTACATGTTGATAAGAACTATCTTACACCA 2301
Db 2221 GAATATAATAAAATATATTCGTTATTATCTACATGTTGATAAGAACTATCTTACACCA 2280
Qy 2302 GACATCAAAAAGAAATATTCGCTTCTATCATTAAGCAGCAAGTGAATATTATTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTCGCTTCTATCATTAAGCAGCAAGTGAATATTATTACTAAAT 2340
Qy 2362 AATGACATCTCATATTACACAGTAATAGACTAATAAAGCTGAGGCACATTTAAGTAAT 2421
Db 2341 AATGATATCTCATATTACACAGTAATAGACTAATAAAGCTGAGGCACATTTAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCATTTTGTGATAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCATTTTGTGATAATCATGAC 2460
Qy 2482 AGCCTATTCGTTAAATATGACAGTATGCTTTATGAAAAATATGATGTCGGCATGAAT 2541
Db 2461 AGCCTATTCGTTAAATATGACAGTATGCTTTATGAAAAATATGATGTCGGCATGAAT 2520
Qy 2542 TTCCTCAGCATTAACACATGATTTGATCGAGAAATCAATGCGCATCCACCATTTTAAAAAG 2601
Db 2521 TTCCTCAGCATTAACACATGATTTGATCGAGAAATCAATGCGCATCCACCATTTTAAAAAG 2580
Qy 2602 CTGATTTAAACCTATTTTAAATGACATGACTTAAGAGATGATGAATGTGAAAGGGGCATCA 2661
Db 2581 CTCATTTAAACCTATTTTAAATGACATGACTTAAGAGATGATGAATGTGAAAGGGGCATCA 2640
Qy 2662 CAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACCATTTATTAAGAAATGTC 2721
Db 2641 CAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTTCTGACCATTTATTAAGAAATGTC 2700
Qy 2722 ATCAGATCTGCCAGTCAATTTGATGAGTGTGCCAGAAATAACACTGAGGATATTTGGTTC 2781
Db 2701 ATCAGATCTGCCAGTCAATTTGATGAGTGTGCCAGAAATAACACTGAGGATATTTGGTTC 2760
Qy 2782 CAATTTGCACCTTTTAAATCTTAGAAAAAGAAACCGGCCATGTTATTTAATAAAACATCGACC 2841

Db 2761 CAATTTGCACCTTTTAAATCTTAGAAAAAGAAACCGCCATGTTATTTAATAAAACATCGACC 2820
Qy 2842 CTGACTTTATATGCTTTGGGAACGAAATTTACAATGGACAAATGAACAAATTTCAAAGTGCA 2901
Db 2821 CTGACTTTATATGCTTTGGGAACGAAATTTACAATGGACAAATGAACAAATTTCAAAGTGCA 2880
Qy 2902 AAAAAAGCGGAAATATCCCGTTTAAACAGTTTCATTTAATAGTATAACGCTATAA 2958
Db 2881 AAAGAGGAGAAATATATACCTGTTAACAGTTTCATTTAATAGTATAACTCTATAA 2937
RESULT 11
PCT-US99-26501-2
; Sequence 2, Application PC/TUS9926501
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 617481-5
; CURRENT APPLICATION NUMBER: PCT/US99/26501
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,929
; EARLIER FILING DATE: 1998-11-11
; EARLIER APPLICATION NUMBER: 09/283,402
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Pasteurella multocida
PCT-US99-26501-2
Query Match 50.5%; Score 1503.2; DB 1; Length 2112;
Best Local Similarity 82.5%; Pred. No. 2.9e-273;
Matches 1743; Conservative 0; Mismatches 348; Indels 21; Gaps 1;
Qy 61 ATGAATACATTATCAAGCAATAAAGCATATACAGCAATGACTATGAAATTAGCACTC 120
Db 1 ATGAATACATTATCAAGCAATAAAGCATATACAGCAATGACTATCAATTAGCACTC 60
Qy 121 AAATTTATTTGGAAGTCTGCTGAAACCTACGGGGGAAAAAATCGTTGAAATTCGAAATATAC 180
Db 61 AAATTTATTTGGAAGTCTGCTGAAACCTATGAGCGGAAAAATCTATGAGCGGAAAAATTTGTTGAAATTTCAAATATAC 120
Qy 181 AAATCTAAGAAAAACTCTCGACCAATTC-----TTATGTAACCT 219
Db 121 AAATGCCAAGAAAAACTCTCAGCACATCTCTGTTAATTCAGCACATCTTTCTGTAAAT 180
Qy 220 GAAGATAAAAAAAGACAGTGTTCGCGATAGCTCATTTAGATATCGCAACACAGCTCTTACTT 279
Db 181 AAGAAGAAAAAAGTCAATGTTTCGCGATAGCTAGTTCGCGATAGATATTCGCAACACAGCTTACTT 240
Qy 280 TCCAAAGTAAAAAATTAACCTCTATCCGAATCAGAAAAAAGAGTTTAAAAAATAAATGG 339
Db 241 TCCAAAGTAAAAAATTAAGTACTTCTGACTCGGAAAAAAGACGTTAAAAAATAAATGG 300
Qy 340 AAATCTATCACTGGGAAAAAATCGGAGACGAGAAATCAGAAAGGTGAAAGCTAGTACCC 399
Db 301 AAATGTCACCTGAGAGAAAAATCTGAAAAATCGGAGGTAAAGCGGTGCCCTTGTACCA 360
Qy 400 AAAGATTTTCTTAAAGATCTTCTGCTTCATTCGATTCGCGATGATCATGTTAATGATTTTACA 459
Db 361 AAAGATTTTCCCAAGATCTGCTTTTAGCGCTTTTACCTGATCATGTTAATGATTTTACA 420
Qy 460 TGGTACAAAAATCGAAAAAAGAGTGTAGGTATAAGCCCTGTGTAATAGAAATATCGGCTTT 519
Db 421 TGGTACAAAAATCGAAAAAAGAGTGTAGGTATAAGCCCTGTGTAATAGAAATATCGGCTTT 480
Qy 520 TCTATTATTTCTTACATTTTAAATCGTACCGCTATTTAGATATATACGTTAGCCCTGTTG 579
Db 481 TCTATTATCGTTACACATTCATCAATCGACAGCAATTTTATCATGATTACATTCGCTGTTTA 540

QY 580 GTCAATCAGAAACAACTACCCATTGGAAGTCGTGTGTCAGATGATGGTAGTAGGAA 639
DB 541 GTAAACCAAAACACATTAACCGTTTGAAGTTATCGTCACAGATGATGGTAGTAGGAA 600
QY 640 AACTTACTTACCATTGTGCAAAATACGAAACAACTTGCATATAAGTAGTAGACACAA 699
DB 601 GATCTATCATCCGATCATTCGCCAATATGAATAAATTCGATATTCGCTACGTCACACAA 660
QY 700 AAGATTTATGGATATCAATTTGTGTCAGTCAGAACTTAGGTTTACGTACAGCAAAAGTAT 759
DB 661 AAGATAACGGTTTCAAGCCAGTCGCGCTCGGAATATGGGATTTAGCGTTAGCAAAATAT 720
QY 760 GATTTTGTCTCGATTCTAGACTCGCATATGGCACCAACAATATATGGGTTTCAATCTTAT 819
DB 721 GACTTTATGGCTTACTCGACTGTGATATGGCGCAAAATCCATTTATGGGTTTCAATCTTAT 780
QY 820 CTTACAGACTATTAGAAGCAATATGATTTGTTTAAATTTGGACCTAGAAAATATGTTGGAT 879
DB 781 GTTGCAGAGCTATTAGAAGATGATGATTTAAACAATCATTTGGTCCAAAGAAATACATCGAT 840
QY 880 ACTCAATAATATACCGCAGAACAAATTCCTTAACGATCCATATTTAATAGAATCACTACCT 939
DB 841 ACACACATATTCACCAAAAGACTTCTTAATAACCGCGATTTGGCTTGAATCATTTACCA 900
QY 940 GAAACCGCTACAATAACCAATCCTCGATTPACATCAAAAAGGAAATATATCGTTGGATGG 999
DB 901 GAAAGTGAACCAACCAATAATAGTTGTCCGCAAAAGGGAAGGAAACAGTTTCTCTGGATTG 960
QY 1000 AGATTAGACATTTCAAAAACCGATATCTACGTCATGTCATTTCTCGTTTTCGTTAT 1059
DB 961 CGCTTAGAACAAATTCGAAAACAGAAAATCTCCGCTTATCCGATTCGCGCTTTCCGTTTT 1020
QY 1060 TTTGTTGCGGGTAATGTTGCTATTTCTAAAGAATGGCTAAATGAAGTAGTGTGGTTCGAT 1119
DB 1021 TTTGCGCGGGTAATGTTGCTTTTCGTAATAAATGGCTAAATAAATCCGTTTCTTTGAT 1080
QY 1120 GAAGAAATTTAATCATTTGGGGGCGGAAGATGTAGAATTTGGTTTACAGATTAATTTGCCAAA 1179
DB 1081 GAGGAATTTAATCACTGGGTGGGAAGATGTGGAATTTGGATATCGCTATTATCCCGTTAC 1140
QY 1180 GCGTGTTTTTCAGAGTAATTTGACGGGGAATGGCCATCCATCAAGAACACCTGGTAAA 1239
DB 1141 GGTAGTTTCTTTAAAACATTTGATGGCTATTTGGCTTACCATCAAGAGCCACCGGTAAA 1200
QY 1240 GAAATGAAACAGACGAGCTGTGTAAGTATTACGCTTAAATTTGTAAGGAAAG 1299
DB 1201 GAAATGAAACCCGATCGTGAAGCGGGAATAATATACGCTCATATATGAGAGAAAG 1260
QY 1300 GTACCTTACATCATAGAAAGCTTTTACCATAGAGATTCACATATTCATAGAATACCT 1359
DB 1261 GTCCCTTATATCATAGAAAACCTTTTACCAATAGAGATTCGATATCAATAGAGTACCT 1320
QY 1360 TTAGTTTCTATTTATATCCCGCTTAACTGTGCAATTTATATCAAGATGTGTAGAT 1419
DB 1321 TTAGTTTCAATTTATATCCCGCTTAACTGTGCAATTTATATCAAGCTTGGTAGAT 1380
QY 1420 AGTCTCTTAATCAAACTGTGTGATCTCGAGTTTGTATTTGTAACGATGTTCACAA 1479
DB 1381 AGTGCACCTGAATCAGACTGTGTTGATCTCGAGTTTGTATTTGTAACGATGTTCACAA 1440
QY 1480 GATAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATATCTAGGTACGCATCATG 1539
DB 1441 GATAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATATCTAGGTACGCATCATG 1500
QY 1540 TCTAAACCAATGCGGAATAGCTCAGCATCAAAATCAGCCGTTTCTTTTGTCTAAAGGT 1599
DB 1501 TCTAAACCAATGCGGAATAGCTCAGCATCAAAATCAGCCGTTTCTTTTGTCTAAAGGT 1560
QY 1600 TATTACATTTGGCAGTTTAGATTCAGATGATTTCTTGGCCTGTAGTGAAGCTGTGT 1659
DB 1561 TATTACATTTGGCAGTTTAGATTCAGATGATTTCTTGGCCTGTAGTGAAGCTGTGT 1620
QY 1660 TTAAGAAGATTTTAAAGATAAAGACCGTAGCTGTGTTTATACCACTAATAGAACGTC 1719

DB 1621 TTAAGAAGATTTTAAAGATAAAGCTAGCTTGTGTTTATACCACTAATAGAAAGCTC 1680
QY 1720 AATCCGGATGGTAGCTTAATCGCTAATGTTACAATTTGGCCAGAAATTTTCCAGAGAAAA 1779
DB 1681 AATCCGGATGGTAGCTTAATCGCTAATGTTACAATTTGGCCAGAAATTTTCCAGAGAAAA 1740
QY 1780 CTCACACCGCTATGATTTGCTCACCATTTTTAGAATGTTTACGATTTAGAGCTTGGCATTTA 1839
DB 1741 CTCACACCGCTATGATTTGCTCACCATTTTTAGAATGTTTACGATTTAGAGCTTGGCATTTA 1800
QY 1840 ACGGATGGATTTAACGAAATATTTGAACCGCGTGGATTTACATGTTTCCCTTAAACTC 1899
DB 1801 ACTGATGGATTTCAATGAAGAAATTTGAATGCGGTAGACTATGACATGTTTCCCTCAAACTC 1860
QY 1900 AGTGAAGTTTGGAAAAATTTAAACATCTTAATAAATCTGCTATAACCCGCTATTACATGGT 1959
DB 1861 AGTGAAGTTTGGAAAAATTTAAACATCTTAATAAATCTGCTATAACCCGCTATTACATGGT 1920
QY 1960 GATAACACATCTTAAAGAACTCGGCAATTTCAAAAGAAAAACCAATTTTGTGTAGTCAAT 2019
DB 1921 GATAACACATCAATTAAGAACTTGGCAATTTCAAAAGAAAAACCAATTTTGTGTAGTCAAT 1980
QY 2020 CAGTCATTAATAGACAAAGCATCAATTAATAATATGACAAATTTGATGATTAGAT 2079
DB 1981 CAGTCATTAATAGACAAAGCATCAATTAATAATATGACAAATTTGATGATTAGAT 2040
QY 2080 GAAAGTAGAAGTATATCTTCAATAAAGCCGCTGAATATCAAGAAAGAAATGGATATGTTA 2139
DB 2041 GAAAGTAGAAGTATATTTTCAATAAAGCCGCTGAATATCAAGAAAGAGATTTGATATCTTA 2100
QY 2140 AAAGATCTTAAA 2151
DB 2101 AAAGATATTAAA 2112

RESULT 12
US-09-107-532-2527
: Sequence 2527, Application US/09107532
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
: TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7308
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD/ROM ISO9660
: COMPUTER:
: OPERATING SYSTEM:
: SOFTWARE:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/ 085598
: FILING DATE: May 14, 1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Ariniello, Pamela Deneke
: REGISTRATION NUMBER: 40,489
: REFERENCE/DOCKET NUMBER: GTC-012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781)893-5007
: TELEFAX: (781)893-8277

```

; INFORMATION FOR SEQ ID NO: 2527:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 984 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: circular
;     MOLECULE TYPE: DNA (genomic)
;     HYPOTHETICAL: NO
;     ANTI-SENSE: NO
;     ORIGINAL SOURCE:
;     ORGANISM: Enterococcus faecalis
;     FEATURE:
;         NAME/KEY: misc_feature
;         LOCATION: 1..984
; US-09-107-532-2527

```

Query Match	3.08;	Score 89.2;	DB 15;	Length 984;
Best Local Similarity	54.2%;	Pred. No. 1.3e-06;		
Matches 181;	Conservative 0;	Mismatches 153;	Indels 0;	Gaps 0;

Qy	1349	ATAGAATACCTTTAGTTTCATATTTATATCCCGCTTATAACTGTCCAAATATATTATCAAA	1408
Db	14	ATAAATGTGTGAGATTAGTATTATTTGTCGTGATATAAAAGTTGAAATAATTTAAGAA	73
Qy	1409	GATCTGTAGATAGTGCCTCTTAATCAAACTGTTGCGATCTCGAGGTTTGTATTTCGTAACG	1468
Db	74	ANTGTGTAGATTCCGATTTTACCCCAACATTTACTGATTTTGAAGTTATATAGTAGATG	133
Qy	1469	ATGTTTCAACAGATAANTACCTTAGAAGTGATCAATAAGCTTTATGGTAAATAATCCTAGGG	1528
Db	134	ATGGATCTCCTGCAATAGCGGGAATACTCTGTGATGAGTATGCTGAAAAGAGATAATCGAG	193
Qy	1529	TACGCATCATGCTTAACCCAAATGGCGGAATAGCCTCAGCATCAAAATCGACCGCTTTCTT	1588
Db	194	TACGTGTGATCCATAAAGAAACCGCGGATTAACGAGTCTAGAAATAACGCGGGATTTGATG	253
Qy	1589	TTGCTTAAAGGTTATTACATTGGCGAGTTAGATTACAGATGATTCTTTCAGCCCTGATCGAC	1648
Db	254	TTGGAGAGGCAATACTTAGTCCTTGTGACAGTGATGATTATATTGATGNAGATATGT	313
Qy	1649	TTGAAGTGTGTTTTAAAGAAATTTTTTAAAGATAA	1682
Db	314	ATGAAATTTTGTATGAGAAATTTGAAATTCATGA	347

RESULT 13
US-09-107-532A-3527
; Sequence 2571 Application US/09107532A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997

```

: ATTORNEY/AGENT INFORMATION:
:
:   NAME: Arintello, Pamela Denette
:   REGISTRATION NUMBER: 40,489
:   REFERENCE/DOCKET NUMBER: GTC-012
:
: TELECOMMUNICATION INFORMATION:
:
:   TELEPHONE: (781)893-5007
:   TELEFAX: (781)893-8277
:
: INFORMATION FOR SEQ ID NO: 2527:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH: 984 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: double
:     TOPOLOGY: circular
:
:   MOLECULE TYPE: DNA (genomic)
:
:   HYPOTHETICAL: NO
:   ANTI-SENSE: NO
:
:   ORIGINAL SOURCE:
:
:     ORGANISM: Enterococcus faecium
:
:   FEATURE:
:
:     NAME/KEY: misc_feature
:     LOCATION: (b) LOCATION 1...984
:     SEQUENCE DESCRIPTION: SEQ ID NO: 2527
:
: US-09-107-532A-2527

```

	Query Match	3.08;	Score 89.2;	DB 15;	Length 984;
	Best Local Similarity	54.29;	Pred. No. 1.3e-06;		
	Matches 181;	Conservative 0;	Mismatches 153;	Indels 0;	Gaps 0;
Qy	1349	ATAGAATACCTTTAGTTTCTATTTATATACCCGCTTATAA	CTGTCAATATATATTTCAA	1408	
Db	14	ATAATAATGTGAGATTAGTATTATTGTTCTGTGTATAA	AGTTGAAAATAATTTAAGAA	73	
Qy	1409	GATGTGTAGATAGTCTCTTAATCAAACTGTTGTGATCT	CGAGGTTTGTATTTCGTAACG	1468	
Db	74	AATGTGTAGATTCGATTTTAGCCCAACATTTTACTGAT	TTTGAACCTATATTAGTAGATG	133	
Qy	1459	ATGGTTCAACAGATAAATACCTTAGAAGTGTATCAATA	AGCTTTATCGTAAATCCTAGGG	1528	
Db	134	ATGGATCTCCCTGACATAAGCGGAAATCTGTGTAGAT	GCTGTGAAAAGATAATACGAG	193	
Qy	1529	TACGCATCATGTCATAAACCAATGCGGAATAGCCTCA	GCAGTCAAAATGCAGCCGTTCTT	1588	
Db	194	TACGTGTGATCCATAAAGAAACGCGGATTTAAGCAG	TGCTTGAAGTTCGGGGATTGATG	253	
Qy	1589	TTGCTTAAAGGTTATTACATTTGGCGAGTTAGATTTC	AGATGATTATCTTGAGCCGTGATGCAG	1648	
Db	254	TGCGAGAGGCAATACCTTAGGCTTTGTTGACAGTG	ATGATTATTTGATGAAGATATGT	313	
Qy	1649	TTGAACCTGTGTTTAAAGAAATTTTTTAAAGATAAA		1682	
Db	314	ATGAAATTTTGTATGAGAAATTTGAAAATTCATGA		347	

```

RESULT 14
US-09-134-000-2987
; Sequence 2987, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: FAECALIS FOR DIAGNOSTI
; FILE REFERENCE: GPC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 2987
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000-2987

```

Query Match 2.9%; Score 86; DB 15; Length 993;
Best Local Similarity 54.0%; Pred. NO. 5.3e-06;
Matches 176; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:32:29 ; Search time 249 seconds
(without alignments)
9950.512 Million cell updates/sec

Title: US-09-842-484A-1

Perfect score: 2979

Sequence: 1 ttataactgattaaagaag.....acattgcatattattataaa 2979

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 657839 seqs, 415857249 residues

Total number of hits satisfying chosen parameters: 1315678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2979	100.0	2979	5	US-09-842-484A-1
2	2963	99.5	2979	5	US-09-842-484A-3
3	2285	76.7	2937	5	US-09-469-2000-8
4	2283.4	76.6	2937	6	US-10-309-560-7
5	2271.2	76.2	2920	6	US-10-217-613-2
6	814.2	27.3	1483	6	US-10-216-289-3
7	807.2	27.1	2058	6	US-10-216-289-1
8	86	2.9	993	5	US-09-134-000C-2987
c 9	79.2	2.7	11474	6	US-10-303-161-1
c 10	79.2	2.7	11474	6	US-10-303-162-1
c 11	79.2	2.7	11474	6	US-10-303-118-1
c 12	79.2	2.7	11474	6	US-10-303-128-1
c 13	79.2	2.7	11474	6	US-10-303-134-1
c 14	77.2	2.6	17527	6	US-10-240-454-27
c 15	76.6	2.6	13825	6	US-10-264-213-27
c 16	74.2	2.5	17276	5	US-09-751-708A-83
17	72.6	2.4	2529	5	US-09-134-000C-2344
18	72.4	2.4	25020	6	US-10-192-280-1
c 19	71.8	2.4	9106	6	US-10-240-452-69
c 20	70	2.3	14924	6	US-10-240-452-21
c 21	69.8	2.3	16258	6	US-10-257-166-120
c 22	68.4	2.3	891	6	US-10-303-161-30
23	68.4	2.3	891	6	US-10-303-162-30
24	68.4	2.3	891	6	US-10-303-118-30
25	68.4	2.3	891	6	US-10-303-128-30
26	68.4	2.3	891	6	US-10-303-134-30

c 27	67.8	2.3	18585	6	US-10-240-485-162	Sequence 162, Appl
28	67.2	2.3	912	6	US-10-303-161-28	Sequence 28, Appl
29	67.2	2.3	912	6	US-10-303-162-28	Sequence 28, Appl
30	67.2	2.3	912	6	US-10-303-118-28	Sequence 28, Appl
31	67.2	2.3	912	6	US-10-303-128-28	Sequence 28, Appl
32	67.2	2.3	912	6	US-10-303-134-28	Sequence 28, Appl
33	66.6	2.2	906	6	US-10-303-161-26	Sequence 26, Appl
34	66.6	2.2	906	6	US-10-303-162-26	Sequence 26, Appl
35	66.6	2.2	906	6	US-10-303-118-26	Sequence 26, Appl
36	66.6	2.2	906	6	US-10-303-128-26	Sequence 26, Appl
37	66.6	2.2	906	6	US-10-303-134-26	Sequence 26, Appl
c 38	66.6	2.2	11836	6	US-10-240-453-114	Sequence 114, Appl
c 39	65.6	2.2	11422	6	US-10-257-166-18	Sequence 18, Appl
c 40	64.2	2.2	7615	6	US-10-257-166-53	Sequence 53, Appl
c 41	64	2.1	14551	6	US-10-240-485-138	Sequence 138, Appl
c 42	63.8	2.1	6106	6	US-10-257-166-113	Sequence 113, Appl
c 43	63.4	2.1	7057	6	US-10-240-485-148	Sequence 148, Appl
c 44	63.4	2.1	20486	6	US-10-240-485-163	Sequence 163, Appl
c 45	63	2.1	6641	6	US-10-240-452-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-09-842-484A-1

; Sequence 1, Application US/09842484A

; GENERAL INFORMATION:

; APPLICANT: DEANGELIS, PAUL L.

; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAME

; FILE REFERENCE: 4605.003

; CURRENT APPLICATION NUMBER: US/09/842.484A

; CURRENT FILING DATE: 2001-04-25

; PRIOR APPLICATION NUMBER: 60/199,538

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 2979

; TYPE: DNA

; ORGANISM: Pasteurella multocida

US-09-842-484A-1

Query Match	100.0%	Score 2979;	DB 5;	Length 2979;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2979;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy 1	TTATAACTGATTAAAGAGGTAAACGATTCAAGCAAGGTTAATTTTAAAGGAAAGAAA	60		
Db 1	TTATAACTGATTAAAGAGGTAAACGATTCAAGCAAGGTTAATTTTAAAGGAAAGAAA	60		
Qy 61	ATGAATACATTATCAACAGCAATAAAGCATATACAGCAATGACTATCATTAATGACACTC	120		
Db 61	ATGAATACATTATCAACAGCAATAAAGCATATACAGCAATGACTATCATTAATGACACTC	120		
Qy 121	AAATTTATTGAGAACTCTGCTGAAACCTACGGCGGAGAAAATCGTTGAAATTCACAAATATC	180		
Db 121	AAATTTATTGAGAACTCTGCTGAAACCTACGGCGGAGAAAATCGTTGAAATTCACAAATATC	180		
Qy 181	AAATTTAAAGAAAACTCTCGACCAATTTTATGTAAGTGAAGATAAAAAACAGTGT	240		
Db 181	AAATTTAAAGAAAACTCTCGACCAATTTTATGTAAGTGAAGATAAAAAACAGTGT	240		
Qy 241	TCCGATAGCTCATTTAGATATCCCAACACAGCTCTTACTTCCCAACGTAATAAATTAAC	300		
Db 241	TCCGATAGCTCATTTAGATATCCCAACACAGCTCTTACTTCCCAACGTAATAAATTAAC	300		
Qy 301	CTATCCGAATCAGAAAAAACAGTTTAAAAATAAATGAAATCTATCACTCGGAAAAAA	360		
Db 301	CTATCCGAATCAGAAAAAACAGTTTAAAAATAAATGAAATCTATCACTCGGAAAAAA	360		
Qy 361	TCCGAGAACGCAAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTTAAAGATCT	420		
Db 361	TCCGAGAACGCAAAATCAGAAAGGTGGAACCTAGTACCCAAAGATTTTCTTAAAGATCT	420		

Db 361 TCGGAGACGCGAAGTTCAGAAAGCTGGNACTAGTACCCAGAGATTTTCCTAAAGATCTT 420
Qy 421 GTTCTTGCTCCATGCCAGATCATGTTAATGATTTTACATGGTACAAAATCGAAAAAA 480
Db 421 GTTCTTGCTCCATGCCAGATCATGTTAATGATTTTACATGGTACAAAATCGAAAAAA 480
Qy 481 AGCTTAGGTATAAAGCCTGTAATAGATATCGGTCTTCTTATTTATTTCTCTCATTTT 540
Db 481 AGCTTAGGTATAAAGCCTGTAATAGATATCGGTCTTCTTATTTATTTCTCTCATTTT 540
Qy 541 AATCGTAGCGGTATTTAGATATAAGCTTAGCTGTTGGTCAATCAGAAAAACAACATAC 600
Db 541 AATCGTAGCGGTATTTAGATATAAGCTTAGCTGTTGGTCAATCAGAAAAACAACATAC 600
Qy 601 CCATTTGAAGTCTGTTTGCAGATGATGTAGTAAGGAAACCTTACTTACCATTGTGCAA 660
Db 601 CCATTTGAAGTCTGTTTGCAGATGATGTAGTAAGGAAACCTTACTTACCATTGTGCAA 660
Qy 661 AATAGCAACAAAACCTTCACATAAAGTATGTAAGCAAAAAAGATTATGGATATCAATTG 720
Db 661 AATAGCAACAAAACCTTCACATAAAGTATGTAAGCAAAAAAGATTATGGATATCAATTG 720
Qy 721 TGTGAGTCAGAACTTAGTTTACGTACAGCAAAAGTATGATTTGCTCGATTCAGAC 780
Db 721 TGTGAGTCAGAACTTAGTTTACGTACAGCAAAAGTATGATTTGCTCGATTCAGAC 780
Qy 781 TGGGATATGCCACCAACAAATTTATGGTTCATCTTACAGAACTTATAGAGAC 840
Db 781 TGGGATATGCCACCAACAAATTTATGGTTCATCTTACAGAACTTATAGAGAC 840
Qy 841 AATGATATGTTTAAATTTGACCTAGAAAATATGTTGGATCTCATATATATCCGCAGAA 900
Db 841 AATGATATGTTTAAATTTGACCTAGAAAATATGTTGGATCTCATATATATCCGCAGAA 900
Qy 901 CAATTCCTTAACGATCCATTTAATAGATCACTACTCTGAAACCGCTACAAATACAAT 960
Db 901 CAATTCCTTAACGATCCATTTAATAGATCACTACTCTGAAACCGCTACAAATACAAT 960
Qy 961 CTTTCGATTAACATCAAAAGGAAATATATCGTTGGATGAGATAGAACATTTCAAAAA 1020
Db 961 CTTTCGATTAACATCAAAAGGAAATATATCGTTGGATGAGATAGAACATTTCAAAAA 1020
Qy 1021 ACCGATTAATCTAGTCTATGATCTCCGTTTCGTTTATTTGTTGGGTAATGTTGCA 1080
Db 1021 ACCGATTAATCTAGTCTATGATCTCCGTTTCGTTTATTTGTTGGGTAATGTTGCA 1080
Qy 1081 TTTTCTAAAGATGGCTAAATAAGTAGTTTGGTTCGATGAAGAAATTAATCATTTGGGG 1140
Db 1081 TTTTCTAAAGATGGCTAAATAAGTAGTTTGGTTCGATGAAGAAATTAATCATTTGGGG 1140
Qy 1141 GCGAAGATGTAGAAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
Db 1141 GCGAAGATGTAGAAATTTGGTTACAGATTAATTTGCCAAAGGCTGTTTTTCAGAGTAAT 1200
Qy 1201 GACGGCGAATGGCCATCCATCAGAACCCACCTGGTGAAGAAATGAACAGACGCGAA 1260
Db 1201 GACGGCGAATGGCCATCCATCAGAACCCACCTGGTGAAGAAATGAACAGACGCGAA 1260
Qy 1261 GCTGTAAGATGATTTACGCTTTAAATTTGTAAGAAAGGTACCTTACATCTATAGAAG 1320
Db 1261 GCTGTAAGATGATTTACGCTTTAAATTTGTAAGAAAGGTACCTTACATCTATAGAAG 1320
Qy 1321 CTTTACCATAAGAGATTCACATATTCATAGATAACCTTTAGTTTCTTATTTATATCCCC 1380
Db 1321 CTTTACCATAAGAGATTCACATATTCATAGATAACCTTTAGTTTCTTATTTATATCCCC 1380
Qy 1381 GCTTATTAAGTGTGAAATTTATTCAGAGTGTAGATGCTCTTAATCAAACTGTT 1440
Db 1381 GCTTATTAAGTGTGAAATTTATTCAGAGTGTAGATGCTCTTAATCAAACTGTT 1440
Qy 1441 GTCGATCTCGAGGTTGTTGTTGTAACGATGTTCAACAGATAATFACCTTAGAAGTGATC 1500
Db 1441 GTCGATCTCGAGGTTGTTGTTGTAACGATGTTCAACAGATAATFACCTTAGAAGTGATC 1500

Qy 1501 AATAAGCTTTATGGTAATAATCCTAGGGTACGCATCATGTCTAAACCAATGCGGAATA 1560
Db 1501 AATAAGCTTTATGGTAATAATCCTAGGGTACGCATCATGTCTAAACCAATGCGGAATA 1560
Qy 1561 GCTCAGCATCAAAATCGACGCGTTCTTTTCTGCTAAAGGTTATTACATTGGCGAGTTAGAT 1620
Db 1561 GCTCAGCATCAAAATCGACGCGTTCTTTTCTGCTAAAGGTTATTACATTGGCGAGTTAGAT 1620
Qy 1621 TCAGATGATTATCTTCAGGCTGATGCAGTTGAACCTGTTGTTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATTATCTTCAGGCTGATGCAGTTGAACCTGTTGTTTAAAGAAATTTTAAAGAT 1680
Qy 1681 AAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTCAATCCGATGCTAGCTTAATC 1740
Db 1681 AAAACGCTAGCTTGTGTTTATACCACTAATAGAAACGTCAATCCGATGCTAGCTTAATC 1740
Qy 1741 GCTAATGTTTACAATTTGGCCAGAAATTTTCACGAGAAAACTCACACGGCTATGATTGCT 1800
Db 1741 GCTAATGTTTACAATTTGGCCAGAAATTTTCACGAGAAAACTCACACGGCTATGATTGCT 1800
Qy 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTAAACGATGATTAAACGAAAT 1860
Db 1801 CACCATTTTGAATGTTTACGATTAGAGCTTGGCATTTAAACGATGATTAAACGAAAT 1860
Qy 1861 ATTGAAACCCGCTGGATTATGACATGTTCCCTTAAACCTCAGTGAAGTTGGAATTTAAA 1920
Db 1861 ATTGAAACCCGCTGGATTATGACATGTTCCCTTAAACCTCAGTGAAGTTGGAATTTAAA 1920
Qy 1921 CATCTTAATAAATCTGCTATACCGGCTATACATGGTGATACACATCCATTAAAGAAA 1980
Db 1921 CATCTTAATAAATCTGCTATACCGGCTATACATGGTGATACACATCCATTAAAGAAA 1980
Qy 1981 CTGGCATTCAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGC 2040
Db 1981 CTGGCATTCAAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGACAGGC 2040
Qy 2041 ATCAATTTATTAATTTATGACAAATTTGATGATTTAGATGAAAGTATATATCTTC 2100
Db 2041 ATCAATTTATTAATTTATGACAAATTTGATGATTTAGATGAAAGTATATATCTTC 2100
Qy 2101 AATAAAACCCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACCTCAAT 2160
Db 2101 AATAAAACCCGCTGAATATCAAGAAAGAAATGGATATGTTTAAAGATCTTAAACCTCAAT 2160
Qy 2161 AATAAAGATGCCAAATCGCGAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Db 2161 AATAAAGATGCCAAATCGCGAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Qy 2221 AAAAACCTTAAACAATATTTATTTGAATATAATAAATAATATTCGTTTATTTACATGTT 2280
Db 2221 AAAAACCTTAAACAATATTTATTTGAATATAATAAATAATATTCGTTTATTTACATGTT 2280
Qy 2281 GATAAGATCATCTTTACACCAGACATCAAAAAGAAATATTGGCTTTCTATCATAGACAC 2340
Db 2281 GATAAGATCATCTTTACACCAGACATCAAAAAGAAATATTGGCTTTCTATCATAGACAC 2340
Qy 2341 CAAGTGAATATTTTACTTAAATAATGACATCTCATATTTACACGAGTAAATAGACTATAAAA 2400
Db 2341 CAAGTGAATATTTTACTTAAATAATGACATCTCATATTTACACGAGTAAATAGACTATAAAA 2400
Qy 2401 ACTGAGGACACATTTAAGTAATATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATAC 2460
Db 2401 ACTGAGGACACATTTAAGTAATATTAATAAATTAAGTCAGTTAAATCTAAATTTGTAATAC 2460
Qy 2461 ATCATTTTGTGATATCATGACAGCCTATTGTTTAAATGACAGCTATGCTTTATATGAAA 2520
Db 2461 ATCATTTTGTGATATCATGACAGCCTATTGTTTAAATGACAGCTATGCTTTATATGAAA 2520
Qy 2521 AATATGATGTCGGCATGAATTTCTCAGATTAAACATGATTGGATCGAGAAAAATCAAT 2580
Db 2521 AATATGATGTCGGCATGAATTTCTCAGATTAAACATGATTGGATCGAGAAAAATCAAT 2580

Qy	2581	CGCGATCCACCATTTTAAAAAGCTGATTAAAAACCTATTATTTTAAATGACAAATGACTTAAAGAAGT	2640
Db	2581	CGCGATCCACCATTTTAAAAAGCTGATTAAAAACCTATTATTTTAAATGACAAATGACTTAAAGAAGT	2640
Qy	2641	ATGAATCTGAAAGGGGCATCACAAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT	2700
Db	2641	ATGAATCTGAAAGGGGCATCACAAAGGTATGTTTATGAAGTATGCGCTACCGCATGAGCTT	2700
Qy	2701	CTGACGATTATTTAAAGAAGTCATCACATCTCTGCGCAATCAATTTGATAGTGTGCCAGAAATAT	2760
Db	2701	CTGACGATTATTTAAAGAAGTCATCACATCTCTGCGCAATCAATTTGATAGTGTGCCAGAAATAT	2760
Qy	2761	AACACTGAGGATATTTGGTTCCCAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCAT	2820
Db	2761	AACACTGAGGATATTTGGTTCCCAATTTGCACCTTTTAACTTTAGAAAAAGAAACCGGCCAT	2820
Qy	2821	GTATTTTAAATAAACATCGACCCCTGACCTTATATGCTTTGGGAACGAAAAATTACAATGGACA	2880
Db	2821	GTATTTTAAATAAACATCGACCCCTGACCTTATATGCTTTGGGAACGAAAAATTACAATGGACA	2880
Qy	2881	AATGAACAAATTTCAAACTGCAAAAAAGCGGAAAAATATCCCGTTTAAACAAGTTCATTATT	2940
Db	2881	AATGAACAAATTTCAAACTGCAAAAAAGCGGAAAAATATCCCGTTTAAACAAGTTCATTATT	2940
Qy	2941	AATAGTATAACGCTATATAACATTTGCAATTTTATATAAA	2979
Db	2941	AATAGTATAACGCTATATAACATTTGCAATTTTATATAAA	2979

RESULT 2
 US-09-842-484A-3
 ; Sequence 3, Application US/09842484A
 ; GENERAL INFORMATION:
 ; APPLICANT: DEANGELIS, PAUL L.
 ; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAME
 ; FILE REFERENCE: 4605.003
 ; CURRENT APPLICATION NUMBER: US/09/842.484A
 ; CURRENT FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/199,538
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 2979
 ; TYPE: DNA
 ; ORGANISM: Pasteurella multocida
 US-09-842-484A-3

Query Match	99.5%	Score	2963;	DB 5;	Length	2979;			
Best Local Similarity	99.7%	Pred. No.	0;						
Matches	2969;	Conservative	0;	Mismatches	10;	Indels	0;	Gaps	0;

Qy	1	TTATAAACTGATTAAAGAAGTTAAACGATTCACGACAGGTTAAATTTTAAAGGAAAGAAA	60
Db	1	TTATAAACTGATTAAAGAAGTTAAACGATTCACGACAGGTTAAATTTTAAAGGAAAGAAA	60
Qy	61	ATGAATACATTATCACAAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC	120
Db	61	ATGAATACATTATCACAAAGCAATAAAAGCATATAACAGCAATGACTATGAATTAGCACTC	120
Qy	121	AAATATTATTGAGAAGTCTGCTGAAACCTACGGCGGAAAAATTCGTTGAATTCCTCAATTTATC	180
Db	121	AAATATTATTGAGAAGTCTGCTGAAACCTACGGCGGAAAAATTCGTTGAATTCCTCAATTTATC	180
Qy	181	AAATGTTAAAGAAAAAATCTCGACCAATTCATGTAAGTGAAGATAAAAAACAGTGT	240
Db	181	AAATGTTAAAGAAAAAATCTCGACCAATTCATGTAAGTGAAGATAAAAAACAGTGT	240
Qy	241	TGCGATAGCTCATTTAGATATCGCAACACAGCTCTTACTTCCCAACGTAATAAAAAATTAAC	300
Db	241	TGCGATAGCTCATTTAGATATCGCAACACAGCTCTTACTTCCCAACGTAATAAAAAATTAAC	300
Qy	301	CTATCCGAATCAGAAAAAACAACAGTTTAAAAAATAATGGAATCTATCACTGGGAAAAA	360

Db	301	 CTATCCGAATCAGAAAAACAGT	 TAAATAAATGGAATCTATC	 ACCTACCTGGGAAAAA	360
Qy	361	 TCGGAGAACGCGAGAAT	 CAGAAGCTGGAAC	 TAGTACCAAGATTTTCT	420
Db	361	 TCGGAGAACGCGAGAAT	 CAGAAGCTGGAAC	 TAGTACCAAGATTTTCT	420
Qy	421	 GTTCTTGCTCCAT	 TGCCAGATCATG	 TAAATGATTTTACATG	480
Db	421	 GTTCTTGCTCCAT	 TGCCAGATCATG	 TAAATGATTTTACATG	480
Qy	481	 AGCTTAGGTATAAGCCT	 GTAAATAGAATATCGG	 TCTTCTATTTATTTCT	540
Db	481	 AGCTTAGGTATAAGCCT	 GTAAATAGAATATCGG	 TCTTCTATTTATTTCT	540
Qy	541	 AATCGTAGCGGTAT	 TTTACATATAACG	 TTTAGCTGTTTGGTCA	600
Db	541	 AATCGTAGCGGTAT	 TTTACATATAACG	 TTTAGCTGTTTGGTCA	600
Qy	601	 CCATTGCAAGTCG	 TGTGCAGATGATG	 TACTAGAGAAAACT	660
Db	601	 CCATTGCAAGTCG	 TGTGCAGATGATG	 TACTAGAGAAAACT	660
Qy	661	 AAATACGACAAAACT	 TGACATAAAGTATG	 TAAGACAAAAAGAT	720
Db	661	 AAATACGACAAAACT	 TGACATAAAGTATG	 TAAGACAAAAAGAT	720
Qy	721	 TGTGTCAGTCAGAACT	 TAGCTACAGCAAAAGT	 ATGATTTGTCCTCCAT	780
Db	721	 TGTGTCAGTCAGAACT	 TAGCTACAGCAAAAGT	 ATGATTTGTCCTCCAT	780
Qy	781	 TGGGATATGGCACCA	 CAAAATATGGGTTC	 ATCTTATCAGAACT	840
Db	781	 TGGGATATGGCACCA	 CAAAATATGGGTTC	 ATCTTATCAGAACT	840
Qy	841	 AATGATATATGTTT	 TAATGGACCTAGAAA	 ATATGCGATCTATA	900
Db	841	 AATGATATATGTTT	 TAATGGACCTAGAAA	 ATATGCGATCTATA	900
Qy	901	 CAATTCCTTAACGAT	 CCATATTAATAGAA	 CTACCTGAACCGCT	960
Db	901	 CAATTCCTTAACGAT	 CCATATTAATAGAA	 CTACCTGAACCGCT	960
Qy	961	 CCTTCGATACATCAA	 AGGAATATCGTT	 GGATGGAGATAGAA	1020
Db	961	 CCTTCGATACATCAA	 AGGAATATCGTT	 GGATGGAGATAGAA	1020
Qy	1021	 ACCGATAATCTACG	 TCTATGTGATCTCG	 TTTCGTTTGGGGT	1080
Db	1021	 ACCGATAATCTACG	 TCTATGTGATCTCG	 TTTCGTTTGGGGT	1080
Qy	1081	 TTTTCTTAAGAAT	 TGCGTAAATAGAGT	 TTTGAATTAATCAT	1140
Db	1081	 TTTTCTTAAGAAT	 TGCGTAAATAGAGT	 TTTGAATTAATCAT	1140
Qy	1141	 GGCGAAGATGAGAA	 TTTGGTTACAGAT	 TATTTGCCAAGGCT	1200
Db	1141	 GGCGAAGATGAGAA	 TTTGGTTACAGAT	 TATTTGCCAAGGCT	1200
Qy	1201	 GACGGCGGAAT	 TGCCATCCAGAAC	 CCCTGGTAAGAAAT	1260
Db	1201	 GACGGCGGAAT	 TGCCATCCAGAAC	 CCCTGGTAAGAAAT	1260
Qy	1261	 GCTGGTAAAGAT	 ATTACGCTTAAAT	 TGGAAGAAAAAGG	1320
Db	1261	 GCTGGTAAAGAT	 ATTACGCTTAAAT	 TGGAAGAAAAAGG	1320
Qy	1321	 CTTTTACCAATAGA	 GATTACATATTC	 ATAGAATACCTTT	1380
Db	1321	 CTTTTACCAATAGA	 GATTACATATTC	 ATAGAATACCTTT	1380
Qy	1381	 GCTTTAACTGTG	 CAAAATATATTC	 AAAGATGTAGAT	1440

Db 1381 GCTTATACCTGCGAAATATATATCAAGATGCTGTAGTAGTGCTCTTAATCAAACTGTT 1440
QY 1441 GTCGATCGAGGTTGTATTTGTAACGATGTTTCAACAGATAATACCTTAGAAGTGATC 1500
Db 1441 GTCGATCGAGGTTGTATTTGTAACGATGTTTCAACAGATAATACCTTAGAAGTGATC 1500
QY 1501 AATAAGCTTTATGTAATAATCCTAGGTAACGATCATCTCTAAACCAATGGCGGAATA 1560
Db 1501 AATAAGCTTTATGTAATAATCCTAGGTAACGATCATCTCTAAACCAATGGCGGAATA 1560
QY 1561 GCGTCAGCATCAAAATGACGCGCTTTCTTTTGTAAAGGTTATACATGCGGAGTTAGAT 1620
Db 1561 GCGTCAGCATCAAAATGACGCGCTTTCTTTTGTAAAGGTTATACATGCGGAGTTAGAT 1620
QY 1621 TCAGATGATTTCTGAGCGCTGATGAGTTGAAGTGTGTTTAAAGAAATTTTAAAGAT 1680
Db 1621 TCAGATGATTTCTGAGCGCTGATGAGTTGAAGTGTGTTTAAAGAAATTTTAAAGAT 1680
QY 1681 AAAACGCTAGCTTGCTTTATACCACTAATAGAAAGCTCAATCCGGATGCTAGCTTTAATC 1740
Db 1681 AAAACGCTAGCTTGCTTTATACCACTAATAGAAAGCTCAATCCGGATGCTAGCTTTAATC 1740
QY 1741 GCTAATGCTTACAAATGCGGAGAAATTTTCAAGAGAAAACCTCACACGCGCTATGATGCT 1800
Db 1741 GCTAATGCTTACAAATGCGGAGAAATTTTCAAGAGAAAACCTCACACGCGCTATGATGCT 1800
QY 1801 CACCAATTTAGATGTTTACGATTTAGAGCTTGCGATTTAACGGATGATTTAACGAAAAT 1860
Db 1801 CACCAATTTAGATGTTTACGATTTAGAGCTTGCGATTTAACGGATGATTTAACGAAAAT 1860
QY 1861 ATTGAAAACCGCGTGGATTTAGCATGTTCCCTTAAACTCAGTGAAGTTGGAATAATTTAAA 1920
Db 1861 ATTGAAAACCGCGTGGATTTAGCATGTTCCCTTAAACTCAGTGAAGTTGGAATAATTTAAA 1920
QY 1921 CATCTTAATTAATCTGCTATATACCGGTATTTACATGCTGATACACATCCATTAAGAAA 1980
Db 1921 CATCTTAATTAATCTGCTATATACCGGTATTTACATGCTGATACACATCCATTAAGAAA 1980
QY 1981 CTCGGCATTCAAAAGAAAACCACTTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGGC 2040
Db 1981 CTCGGCATTCAAAAGAAAACCACTTTTGTGTAGTCAATCAGTCATTTAAATAGACAAGGC 2040
QY 2041 ATCAATTTATTAATTAATGCAAAATTTGATGATTTAGATGAAAGTGAAGTATATCTTC 2100
Db 2041 ATCAATTTATTAATTAATGCAAAATTTGATGATTTAGATGAAAGTGAAGTATATCTTC 2100
QY 2101 AATAAACCCTGATATCAAGAGAAATGATATGTTTAAAGATCTTAAACTCATTCAA 2160
Db 2101 AATAAACCCTGATATCAAGAGAAATGATATGTTTAAAGATCTTAAACTCATTCAA 2160
QY 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
Db 2161 AATAAGATGCAAAATCGCAGTCAGTATTTCTATCCCAATACATTAACGGCTTAGTG 2220
QY 2221 AAAAACTAACAATATTTGATATATAATAAATAATATTCGTTATTTCTACATGTT 2280
Db 2221 AAAAACTAACAATATTTGATATATAATAAATAATATTTCTATCCCAATACATTAACGGCTTAGTG 2280
QY 2281 GATAAGATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATAAGCAC 2340
Db 2281 GATAAGATCATCTTACACAGACATCAAAAAGAAATATTTGGCTTTCTATCATAAGCAC 2340
QY 2341 CAAGTGAATATTTACTAAATTAATGACATCTCATATTTACAGAGTAAATAGACATATAAA 2400
Db 2341 CAAGTGAATATTTACTAAATTAATGACATCTCATATTTACAGAGTAAATAGACATATAAA 2400
QY 2401 ACTGAGGACATTTAAGTAATTAATAAATAAGTCAAGTTAAATCTAAATTTGTAATAC 2460
Db 2401 ACTGAGGACATTTAAGTAATTAATAAATAAGTCAAGTTAAATCTAAATTTGTAATAC 2460
QY 2461 ATCAATTTTGTATATCATCAGACGCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520
Db 2461 ATCAATTTTGTATATCATCAGACGCTATTCGTTAAAAATGACAGCTATGCTTTATATGAAA 2520

QY 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAACACATGATTTGGATCGAGAAAAATCAAT 2580
Db 2521 AAATATGATGTGGCATGAATTTCTCAGCATTAACACATGATTTGGATCGAGAAAAATCAAT 2580
QY 2581 GGGCATCCACCATTTAAAAGCTGATTAACACCTATTTTAAATCACAATGACTTTAAGAAGT 2640
Db 2581 GGGCATCCACCATTTAAAAGCTGATTAACACCTATTTTAAATCACAATGACTTTAAGAAGT 2640
QY 2641 ATGAATGTAAAAGGGGCATCACAAAGTATGTTTATGAAGTATCGCTACCGCATGAGCTT 2700
Db 2641 ATGAATGTAAAAGGGGCATCACAAAGTATGTTTATGAAGTATCGCTACCGCATGAGCTT 2700
QY 2701 CTGACGATTTATTAAGAAAGTCATCATCTGCGCAATCAATTTAGTGTGCGCAGATAT 2760
Db 2701 CTGACGATTTATTAAGAAAGTCATCATCTGCGCAATCAATTTAGTGTGCGCAGATAT 2760
QY 2761 AACACTGAGGATATTTGGTTCCTCAATTTGCACTTTTAAATCTAGAAAAAGAAACCGGCCAT 2820
Db 2761 AACACTGAGGATATTTGGTTCCTCAATTTGCACTTTTAAATCTAGAAAAAGAAACCGGCCAT 2820
QY 2821 GTATTTAATAAATACATCGACCCCTGACTTATATGCTTTGGGAACGAAATTTACAATGGACA 2880
Db 2821 GTATTTAATAAATACATCGACCCCTGACTTATATGCTTTGGGAACGAAATTTACAATGGACA 2880
QY 2881 AATGAACAATTTCAAAGTGCAGAAAAAGGCGAAATATCCCGTTAACAAAGTTTCATTATT 2940
Db 2881 AATGAACAATTTCAAAGTGCAGAAAAAGGCGAAATATCCCGTTAACAAAGTTTCATTATT 2940
QY 2941 AATAGTATAACGCTATATAAATTTGCACTTTTATTAATAA 2979
Db 2941 AATAGTATAACGCTATATAAATTTGCACTTTTATTAATAA 2979

RESULT 3
US-09-469-200D-8
; Sequence 8, Application US/09469200D
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/09/469,200D
; PRIOR FILING DATE: 1999-12-21
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 2937
; TYPE: DNA
; ORGANISM: Pastuerella Multocida
US-09-469-200D-8

Query Match 76.7%; Score 2285; DB 5; Length 2937;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2541; Conservative 0; Mismatches 375; Indels 21; Gaps 1;

QY 43 ATTTTAAAGGAAAGAAATGAATATATCACAAGCAATAAAGCATATTAACAGCAAT 102
Db 1 ATTTTAAAGGACACAGAAATGAATATATCACAAGCAATAAAGCATATTAACAGCAAT 60
QY 103 GACTATGAATTTAGCACTCAAAATTTTGAAGTCTGCTGAACCTACCGGGGAAAAATC 162
Db 61 GACTATCAATTTAGCACTCAAAATTTTGAAGTCTGCTGAACCTACCGGGGAAAAATC 120
QY 163 GTTGAATTTCAAAATTTATCAAAATGTAAGAAAAAATCTCTCGACCAATTC----- 209
Db 121 GTTGAATTTCAAAATTTATCAAAATGTAAGAAAAAATCTCTCGACCAATTCCTCTGTTAATTC 180

Qy	210	-----TTATGTAAGTGAAGATAAAAAACAGTGGTTTGGGATAGCTCAATTAGATATC	261
Db	181		
Qy	262	GCAACACAGCTCTTACTTTCCACAGTAAATAAATAAAGTCTATCCGAATCAGAAAAAAC	321
Db	241		
Qy	322	AGTTTAAAAAATAAATGGAAATCTACTCGGAAAAAATCGGAGAACCGAGAAATCAGA	381
Db	301		
Qy	382	AAGTGGAACTAGTACCCAAAGATTTTCTAAGATCTTGTTCTTGCTTCATGCGCAGAT	441
Db	361		
Qy	442	CATGTTAATCATTTTACATGGTACAAAAATCGAAAAAAGCTTAGGTATATAAGCCCTGA	501
Db	421		
Qy	502	AATAAGAAATACGGTCTTCTTATTTATTTCTACATTTAATCGTAGCGGTATTTTAGAT	561
Db	481		
Qy	562	ATAAGCTTACCGTGTGGTCAATCAGAAAAACAATACCCATTTTGAAGTCGTTGTGGCA	621
Db	541		
Qy	622	GATGATGGTAGTAAGAAAACTTACTTACCATTGTGCAAAAATACGAACAAAAAATTCAC	681
Db	601		
Qy	682	ATAAGTATTCTAAGCAAAAAAGATTATGGATATCAATTTGTGTCAGTCAGAAAACTTAGT	741
Db	661		
Qy	742	TTAGTACAGCAAAAGTATGATTTTGTCTCGATTCTAGACTGCGATATGCCACCACAAAC	801
Db	721		
Qy	802	TTATGGGTTCATTTCTTACAGAACTATTTAGAAGACAATGATATTGTTTAAATTTGGA	861
Db	781		
Qy	862	CCTAGAAAATATGTGGTACTCATATAATTACCAGCAACAATTCCTTTAACGATCCATAT	921
Db	841		
Qy	922	TTAATAGAATCTACTCTGAAACCGGTACAAATAACAATCCTTCGATTACATCAAAAGGA	981
Db	901		
Qy	982	AATATATCGTTGGATTTGGAGATTAGAACATTTCAAAAAACCGGATAATCTACGCTATGT	1041
Db	961		
Qy	1042	GATTCGCGTTTCGTTATTTTGTGGGGTAATTTTGCATTTTCTTAAAGATGGCTAAAT	1101
Db	1021		
Qy	1102	AAAGTAGGTTGGTTCGATGAAGAAATTAATCATTTGGGGGGCGAAGATGTAGAATTTGGT	1161
Db	1081		
Qy	1162	TACAGATTATTTGCCAAAGCTGTTTTTTCAGAGTAATTGACGGCGGAATGGCCATCCAT	1221
Db	1141		
Qy	1222	CAAGAACCCCTGTAAGAAAAATGAACAGAACCGGAAGCTGGTAAAAGTATTACGCTT	1281
Db	1201		
Qy	1282	ANAAATTGAAAGAAAAAGGTACCTTACATCTATAGAAAGCTTTTACCNAATAGAAGATTCA	1341

[illegible]

Qy 1282 AAAATTGTGAAAGAAAGGTACCTTACATCTATAGAAAGCTTTACCAATAGAAGTCA 1341
Db 1261 GATATTATGAGAGAAAGGTCCCTTATATCTATAGAAAACCTTTACCAATAGAAGTTCG 1320
Qy 1342 CATATTATAGAAACCTTTTATGTTTCCTATTATATATCCCGCTTATAAAGTGTGCAAAATAT 1401
Db 1321 CATATCAATAGAGTACCTTTTATGTTTCAATTTTATATCCAGCTTATAAAGTGTGCAAACTAT 1380
Qy 1402 ATTCAAAGATCTGTAGATAGTCTTAAATCAAACTGTGTCGATCTCGAGGTTTGTATT 1461
Db 1381 ATTCAAGCTTGGTAGATAGTCTGAATCAGACTGTGTTGATCTCGAGGTTTGTATT 1440
Qy 1462 TGTAAACGATGGTTCACACAGATAAATACCTTAGAAGTGATCAATAAGCTTTATGTTAAATAT 1521
Db 1441 TGTAAACGATGGTTCACACAGATAAATACCTTAGAAGTGATCAATAAGCTTTATGTTAAATAT 1500
Qy 1522 CCTAGGTTAGCATCATGTCTAAACCAAAATGGCGAATAGCCTCAGCATCAAAATGCGACC 1581
Db 1501 CCTAGGTTAGCATCATGTCTAAACCAAAATGGCGAATAGCCTCAGCATCAAAATGCGACC 1560
Qy 1582 GTTCTTTTGTCTAAAGGTTATTACATTTGGGAGTTAGATTGATGATGATTTCTGAGGCT 1641
Db 1561 GTTCTTTTGTCTAAAGGTTATTACATTTGGGAGTTAGATTGATGATGATTTCTGAGGCT 1620
Qy 1642 GATCAGCTTGAACGTGTGTTTAAAGAAATTTTAAAGATAAAAGCGCTAGCTTTGTGTTTAT 1701
Db 1621 GATCAGCTTGAACGTGTGTTTAAAGAAATTTTAAAGATAAAAGCGCTAGCTTTGTGTTTAT 1680
Qy 1702 ACCACTAATAGAAACGTCAATCCGGATGGTATGCTTAATFCGCTAATGTTTACAAATGGCCA 1761
Db 1681 ACCACTAATAGAAACGTCAATCCGGATGGTATGCTTAATFCGCTAATGTTTACAAATGGCCA 1740
Qy 1762 GAATTTTCAGAGAAAACCTCACACGCTATGATGCTCACCATTTTAGAAATGTTTACG 1821
Db 1741 GAATTTTCAGAGAAAACCTCACACGCTATGATGCTCACCATTTTAGAAATGTTTACG 1800
Qy 1822 ATTAGAGCTTGGCATTTTAAACGGATGGATTTTAAACGAAAATATTTGAAACCGCTGATTTAT 1881
Db 1801 ATTAGAGCTTGGCATTTTAAACGGATGGATTTTAAACGAAAATATTTGAAACCGCTGATTTAT 1860
Qy 1882 GACATGTTCTTAACTCAGTGAAGTTGGAAAATTTAAACATCTTAAATAATCTGCTAT 1941
Db 1861 GACATGTTCTTAACTCAGTGAAGTTGGAAAATTTAAACATCTTAAATAATCTGCTAT 1920
Qy 1942 AACCGCTGATTACATGTTGATTAACACATCCATTAAGAACTCGCATTCACAAAGAAAAC 2001
Db 1921 AACCGTGATTACATGTTGATTAACACATCAATTAAGAACTTGGCATTCACAAAGAAAAC 1980
Qy 2002 CATTTTGTGTPAGTCAATCAGTCAATTAATAGACAAGGCATCAATTTATTAATATATGAC 2061
Db 1981 CATTTTGTGTPAGTCAATCAGTCAATTAATAGACAAGGCATCAATTTATTAATATATGAC 2040
Qy 2062 AAATTTGATGATTTAGATGAAGTAGAAGTATATCTTCAATAAAACCGCTGAATATCAA 2121
Db 2041 GAATTTGATGATTTAGATGAAGTAGAAGTATATTTTCAATAAAACCGCTGAATATCAA 2100
Qy 2122 GAAGAAATGGATGTTTAAAGATCTTAAACTCAATTCAAAATAAAGATGCCAAAATCGCA 2181
Db 2101 GAAGAGATTGATTTTAAAGATATTAAATCATCCAGATAAAGATGCCAAAATCGCA 2160
Qy 2182 GTCAGTATTTTCTATCCCAATPACATTTAAACGGCTTAGTGAAGAAAACCTAAACAAATATTAT 2241
Db 2161 GTCAGTATTTTCTATCCCAATPACATTTAAACGGCTTAGTGAAGAAAACCTAAACAAATATTAT 2220
Qy 2242 GAATATATAAATAATATTCTGTTATTATTTCTACATGTTGATAGAATCATCTTACACCA 2301
Db 2221 GAATATATAAATAATATTCTGTTATTATTTCTACATGTTGATAGAATCATCTTACACCA 2280
Qy 2302 GACATCAAAAAGAAATATTGCTTTCTATCATAGCACCAGTGAATATTTTTACTAAAT 2361
Db 2281 GATATCAAAAAGAAATATTGCTTTCTATCATAGCACCAGTGAATATTTTTACTAAAT 2340

Qy 2362 AATGACATCTCATATTACACGAGTAATAGACTAATAAAAACTGAGGCACATTTTAAGTAAT 2421
Db 2341 AATGATATCTCATATTACACGAGTAATAGACTAATAAAAACTGAGGCAGATTTAAGTAAT 2400
Qy 2422 ATTAATAAATTAAGTCAGTTAAATCTAAATGTGAATACATCATTTTGTATTAATCATGAC 2481
Db 2401 ATTAATAAATTAAGTCAGTTAAATCTAAATGTGAATACATCATTTTGTATTAATCATGAC 2460
Qy 2482 AGCCTATTGTTTAAATAATGACAGCTATGCTTATATGAAAAATATCATCTCGCATGAAT 2541
Db 2461 AGCCTATTGTTTAAATAATGACAGCTATGCTTATATGAAAAATATCATCTCGCATGAAT 2520
Qy 2542 TTCTCAGCATTAACACATGATTGGATCGAGAAAATCAATGCGCATCCACCATTTAAAAAG 2601
Db 2521 TTCTCAGCATTAACACATGATTGGATCGAGAAAATCAATGCGCATCCACCATTTAAAAAG 2580
Qy 2602 CTGATTAAACCTTATTAAATGACAATGACTTAAAGAATGATGAATGTGAAGGGGCATCA 2661
Db 2581 CTGATTAAACCTTATTAAATGACAATGACTTAAAGAATGATGAATGTGAAGGGGCATCA 2640
Qy 2662 CAAGGTATGTTTATCAAGTATGCGCTACCGCATGAGCTTCTGACGATTTTAAAGAAGTC 2721
Db 2641 CAAGGTATGTTTATGACGTATGCGCTAGCGCATGAGCTTCTGACGATTTTAAAGAAGTC 2700
Qy 2722 ATCACATCTCGCAATCAATTTGATAGTGTGCCAGAAATATAACACTGAGGATATTTGGTTC 2781
Db 2701 ATCACATCTCGCAATCAATTTGATAGTGTGCCAGAAATATAACACTGAGGATATTTGGTTC 2760
Qy 2782 CAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCGCATGTTATTAAACACATCGACC 2841
Db 2761 CAATTTGCACTTTTAACTTTAGAAAAGAAACCGGCGCATGTTATTAAACACATCGACC 2820
Qy 2842 CTGACTTATATGCTTGGGACGAAAATTTACAATGGACAAATGAACAAATTTCAAACTGCA 2901
Db 2821 CTGACTTATATGCTTGGGACGAAAATTTACAATGGACAAATGAACAAATTTGAAGTGCA 2880
Qy 2902 AAAAAAGCGAAAATATATCCCGTTAACCAAGTTTCATTATTAAATAGTATAACGCTATAA 2958
Db 2881 AAAAAAGCGAAAATATATATCCCTGTTAAACAAGTTTCATTATTAAATAGTATAACCTATAA 2937

RESULT 5
US-10-217-613-2
; Sequence 2, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AN
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 35341.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-217-613-2

Query Match 76.2%; Score 2271.2; DB 6; Length 2920;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 373; Indels 21; Gaps 1;
Qy 61 ATGAATACATTATCAACAGCAATAAAGCATATACAGCAATGACTATGAATTAGCACTC 120
Db 1 ATGAATACATTATCAACAGCAATAAAGCATATACAGCAATGACTATGAATTAGCACTC 60

QY 121 AAATTTATTTGAGAGTCTGCTGAAACCTACGGCGGAAATAATCGTTGAATTCCAAATATATC 180
DB 61 AAATTTATTTGAAAGTCGGCGGAATCTATGACGGAAATTTGTTGAATTTCAATATACC 120
QY 181 AAATGTAAGAAATACTCTCGACCAATCTTATG-----TAAGT 219
DB 121 AAATGCAAGAAATACTCTCAGCACATCTCTGTTAATTCAGCACATCTTCTGTGTAAT 180
QY 220 GAAGATAAAAAACACAGTGTTCGGATAGCTCATTTAGATATCGCAACACAGCTCTTACAT 279
DB 181 AAAGAAGAAAAAGTCAATGTTTCGGATAGTCGGTTAGATATTTGCAACACAACTGTTTACT 240
QY 280 TCCACGTAATAAAATTAACCTCTATCCGAATCAGAAAAACAGTTTAAAAATAAATG 339
DB 241 TCCACGTAATAAAATTAAGTACTTCTGACTCGGAATAAACACGTTTAAAAATAAATG 300
QY 340 AAATCTATCACTGGGAAAAATCGGAGAACCGCAGAAATCAGAAAGGTGGAACCTAGTACCC 399
DB 301 AAATTCCTCAGAGAAATCTGAAATCGGGAGTAAAGCGCTGCGCCCTGTGACCA 360
QY 400 AAGATTTTCCTAAGATCTTGTCTGTCCATTCGCCAGATCATGTTAATGATTTTACA 459
DB 361 AAGATTTTTCCCAAGATCTGTTTGTAGCGCTTTACCTGATCATGTTAATGATTTTACA 420
QY 460 TGTACAAAAATCGAAAAAAGCTTAGGTATAAAGCCTGTAATAAGAAATATCGGCTT 519
DB 421 TGTACAAAAAGCGAAGAAAGACTTGGCATAAACCTGAACATCAACATGTTGCTT 480
QY 520 TCTATTATTCTCATAATTAATCGTAGCCGTATTTTAGATATTAACGTTAGCTGTTG 579
DB 481 TCTATTATCTTACAACATTCATCGACACCAATTTTATFCGATTTACATAGCTGTTA 540
QY 580 GTCAATCAGAAAAACAACTACCCATTTGAAGTCGTTGTCAGATGATGTTAGTAAAGAA 639
DB 541 GTAAACAAAAACACATTAACCGTTTGAAGTTATCGTGACAGATGATGTCAGGAA 600
QY 640 AACTTACTTACCATTGTCAAAAATACGAAACAAACTTGACATAAAGTATGTAAGCAA 699
DB 601 GATCTATCCCGATTCATTCGCCAATATGAATAAATTTGGATATTCGCTACGTCAGCAA 660
QY 700 AAGATTTATCGATCAATTTGTCGAGTCAGAAACCTTAGTGTACGTACAGCAAGATAT 759
DB 661 AAGATTAACGGTTTCAACGACGTCGCGCTCGGAATATGGGATTTACGTTAGCAAAATAT 720
QY 760 GATTTGTCGATTTAGACTCGGATATGGCACCAACAATATATGGGTTCAATCTTAT 819
DB 721 GACTTTATGCTTACTCGACTGATATGGCGCAATCCATTTATGGGTTCAATCTTAT 780
QY 820 CTTACAGAACTATTAGAAGCAATGATATGTTTTAAATGGACCTAGAAATATGTCGAT 879
DB 781 GTTCGAGAGCTATTAGAAGATGATGATTTAACCAATCATTTGTCACAGAAATATACATCGAT 840
QY 880 ACTCATTAATATACCGCAGAACATTCCTTAACGATCCATATTTAATAGAACTACTACCT 939
DB 841 ACACAACATATTGACCAAAAGACTTCTTAATAACCGGAGTTTGCCTGATCATATTACCA 900
QY 940 GAAACCGCTACAAATAACAATCTCTCGATTACATCAAAAGGAAATATATCGTTGGATGG 999
DB 901 GAAGTGAACCAATAATAGTTTGGCGCAAAAGGGAAGGAAACAGTTTCTCTGGATTGG 960
QY 1000 AGATTAGAACATTTCAAAAACCGAATACTACGTCTATGTGATCTCCGTTTTCGTTAT 1059
DB 961 CGCTTAGAACAAATTCGAAAAAACAGAAATCTCCGCTTATCCGATTCGCCCTTCCGTTT 1020
QY 1060 TTTGTTGCGGTAATGTTGCAATTTCTAAAGAATGGCTTAATAAAGTAGGTTGGTTCGAT 1119
DB 1021 TTTGCGGGGTATGTTGCTTTCGCTAAAAAATGGCTTAATAAATCCGTTTCTTTGAT 1080
QY 1120 GAAGAAATTAATCATTTGGGGGGCGAGATCTAGAAATTTGGTTACAGATTTATTTGCCAAA 1179
DB 1081 GAGGAATTTAATCATCTGGGGTGGAGAAGATGGAATTTGGATATCGCTTATTCGGTTAC 1140
QY 1180 GGCTGTTTTTTCAGAGTAATTTACGGCGGAATGGCCATCCATCAAGAACCCCTGGTAAA 1239

DB 1141 GGTTAGTTCTTTAAAACTATTGATGCAATATGSCCTACCATCAAGGCGCACAGGTAAA 1200
QY 1240 GAAATGAACACGAAACCGGAAAGCTGTAAAGATTATACGCTTAAAAATTGTTGAAGAAAAAG 1299
DB 1201 GAAATGAACACCGATCGTGAAGCGGAAAAAATATTACGCTCGATATTATGAGAGAAAG 1260
QY 1300 GTACCTTACATCTATAGAAGCTTTTACCAATAGAAGATTACATATTCATAGAATACCT 1359
DB 1261 GTCCCTTATATCTATAGAAAACCTTTTACCAATAGAAGATTCGCATATCAATAGATACCT 1320
QY 1360 TTAGTTTCTATTATATPCCCGCTTATACTGTGCAAAATTTATTTCAAGATGTTGTAGAT 1419
DB 1321 TTAGTTTCAATTTATATCCAGCTTATACTGTGCAAACTATATTCAACGTTCCGTTAGAT 1380
QY 1420 AGTGCTTTAATCAAACTGTTGTCGATCTCGAGTTTGTATTTGTAACGATGTTTCAACA 1479
DB 1381 AGTGCACTGAATCAGACTGTTGTTGATCTCGAGTTTGTATTTGTAACGATGTTTCAACA 1440
QY 1480 GATAATACCTTAGAAGTGATCAATAAGCTTTTATGGTAATTAATCCTAGGGTACGCATCATG 1539
DB 1441 GATAATACCTTAGAAGTGATCAATAAGCTTTTATGGTAATTAATCCTAGGGTACGCATCATG 1500
QY 1540 TCTAAACCAATGGCGGAATAGCCTCAGCATCAAAATCGACCGCTTCTTTTCTGCTAAAGGT 1599
DB 1501 TCTAAACCAATGGCGGAATAGCCTCAGCATCAAAATCGACCGCTTCTTTTCTGCTAAAGGT 1560
QY 1600 TATTACATTTGGGCGAGTTAGATTTCAGATGATTATCTTCAGGCTGATGAGTTGAACTGCTGT 1659
DB 1561 TATTACATTTGGGCGAGTTAGATTTCAGATGATTATCTTCAGGCTGATGAGTTGAACTGCTGT 1620
QY 1660 TTTAAAGAAATTTTAAAGATAAACCTAGCTGTTGTTTATACCACATAAGAAAGCTC 1719
DB 1621 TTTAAAGAAATTTTAAAGATAAACCTAGCTGTTGTTTATACCACATAAGAAAGCTC 1680
QY 1720 AATCCGATGCTAGCTTAATCGCTAAATGTTTACAATTTGGCCAGAAATTTTCACAGAAAAA 1779
DB 1681 AATCCGATGCTAGCTTAATCGCTAAATGTTTACAATTTGGCCAGAAATTTTCACAGAAAAA 1740
QY 1780 CTCACAAACGCTATGATTGCTCACCATTTTAGAATGTTTAGGATTTAGGATTTAGGCTTGGCATTTA 1839
DB 1741 CTCACAAACGCTATGATTGCTCACCATTTTAGAATGTTTAGGATTTAGGCTTGGCATTTA 1800
QY 1840 ACGGATGATTTACGAAATATTGAAACCGCTGGATTTAGCATGTTGCTTCTTAAACTC 1899
DB 1801 ACTGATGATTTCAATGAAATAATTTGAAATTTCCGCTAGACTATGACATGTTCTCTCAACTC 1860
QY 1900 AGTGAAGTTGGAATAATTTAAACATCTTAATAAATCTGCTATAACCGCTATTACATGGT 1959
DB 1861 AGTGAAGTTGGAATAATTTAAACATCTTAATAAATCTGCTATAACCGCTGTTTACATGGT 1920
QY 1960 GATAACACATCCATTAAAGAACTCGGCATTTCAAAGAAAAACCATTTTGTGTTAGTCAAT 2019
DB 1921 GATAACACATCAATTAAAGAACTTTGGCATTTCAAAGAAAAACCATTTTGTGTTAGTCAAT 1980
QY 2020 CAGTCATTAAATACAGAGGCTCAATTTATTTATGACAAATTTGATGATTAGAT 2079
DB 1981 CAGTCATTAAATACAGAGGCTCAATTTATTTATGACAAATTTGATGATTAGAT 2040
QY 2080 GAAAGTAGAAAGTATATCTTCAATAAAACCGCTGAATATCAAGAAAGAAATGGATGTTA 2139
DB 2041 GAAAGTAGAAAGTATATTTTCAATAAAACCGCTGAATATCAAGAAAGAGATTGATCTTA 2100
QY 2140 AAGATCTTAAACATTTCAAAATAAAGATCCAAAATCGCAGTCAGATTTCCTATCC 2199
DB 2101 AAGATATTAATAATCATCCAGAATAAAGATCCAAAATCGCAGTCAGATTTCCTATCC 2160
QY 2200 AATACATTTAAACCGCTTAGTGAAAAAATAACAAATATTATTGAATATAATAAATAATA 2259
DB 2161 AATACATTTAAACCGCTTAGTGAAAAAATAACAAATATTATTGAATATAATAAATAATA 2220
QY 2260 TTCGTTATTATCTACATGTTGATAAGAAATCATCTTACACAGACATCAAAAAGAAATA 2319

Db 2221 TTCGTTATTGTTTACATGTTGATAAGAAATCATCTTACACAGATATCAAAAAAGAAATA 2280
Qy 2320 TTGGCTTTCTATCAAGCAACCAAGTGAATATTTTACTAAATAATGACATCTCATATTAC 2379
Db 2281 CTAGCCTTCTATCAACATCAAGTGAATATTTTACTAAATAATGATATCTCATATTAC 2340
Qy 2380 ACGAGTAAATAGACTAAATAAAACATGAGGCACATTTAAGTAAATATTAATAAATTAAGTCAG 2439
Db 2341 ACGAGTAAATAGATTAAATAAAACATGAGGCATTTAAGTAAATATTAATAAATTAAGTCAG 2400
Qy 2440 TTAATCTAAATTTGTAATACATCATTTTTCATATCATGACAGCCTATTCGTTAAATAAT 2499
Db 2401 TTAATCTAAATTTGTAATACATCATTTTTCATATCATGACAGCCTATTCGTTAAATAAT 2460
Qy 2500 GACAGCTATGCTTATATGAAAAAATATGATCGGCATGAATTTCTCAGCATTAACACAT 2559
Db 2461 GACAGCTATGCTTATATGAAAAAATATGATCGGCATGAATTTCTCAGCATTAACACAT 2520
Qy 2560 GATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAGCTGATTTAAAAACCTATTTT 2619
Db 2521 GATTGGATCGAGAAATCAATGCGCATCCACCATTTAAAAAGCTGATTTAAAAACCTATTTT 2580
Qy 2620 AATGACATGACTTAAAGAGTATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAG 2679
Db 2581 AATGACATGACTTAAAGAGTATGAATGTGAAGGGGCATCACAGGTATGTTTATGAAG 2640
Qy 2680 TATGCGCTACCGCATGAGCTTCTGACGATTTATTAAGAGTATCATCACATCTGCGCAATCA 2739
Db 2641 TATGCGCTACCGCATGAGCTTCTGACGATTTATTAAGAGTATCATCACATCTGCGCAATCA 2700
Qy 2740 ATTGATAGTGTCCAGAAATATAACACTGAGGATATTTGGTTCCCAATTTGCACTTTTAATC 2799
Db 2701 ATTGATAGTGTCCAGAAATATAACACTGAGGATATTTGGTTCCCAATTTGCACTTTTAATC 2760
Qy 2800 TTAGAAAGAAACCGGCATGATTTTAAATAAATCAATGACCTGACCTATATGCTTGG 2859
Db 2761 TTAGAAAGAAACCGGCATGATTTTAAATAAATCAATGACCTGACCTATATGCTTGG 2820
Qy 2860 GAACGAAATTTACAATGACAAATGAACAAATTCAAAGTGCAGAAAGGCAAAATATC 2919
Db 2821 GAACGAAATTTACAATGACAAATGAACAAATTTGAAAGTGCAGAAAGGCAAAATATATA 2880
Qy 2920 CCCGTTAAACAGTTCATTTAATAGTATAACGGTATAAAA 2959
Db 2881 CCTGTTAAACAGTTCATTTAATAGTATAACCTCTATAAA 2920

RESULT 6
US-10-216-289-3
; Sequence 3, Application US/10216289
; GENERAL INFORMATION:
; APPLICANT: NINOMIYA, TOSHIO
; APPLICANT: SUGIURA, NOBUO
; APPLICANT: KIMATA, KOJI
; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME
; FILE REFERENCE: 226882USO
; CURRENT APPLICATION NUMBER: US/10/216.289
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: JP 2001-244685
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: JP 2001-324127
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: JP 2002-103136
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 14483
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3787)..(5847)

; OTHER INFORMATION:
US-10-216-289-3
Query Match 27.3%; Score 814.2; DB 6; Length 14483;
Best Local Similarity 63.0%; Pred. No. 1.6e-146;
Matches 1312; Conservative 0; Mismatches 758; Indels 13; Gaps 3;
Qy 53 GAAAGAAAATGAATACATTATCACAAGCAATAAAAGCATATAACAGCAATGACTATGAAT 112
Db 3779 GAAATAATATGATATTCTTAATCAAGCAATAAATTTATATAAAACAAAAATATCGCC 3838
Qy 113 TAGCACTCAAAATATTGAGAAAGTCTGCGAAACCTACGGCGGAAAAATCGTTGAAATCC 172
Db 3839 AAGCTTTATCTCTTTTGGAGAAAGTGTGAAATTTATGATGTTAGTGGTCGGAAGCAA 3898
Qy 173 AAATTTATCAAAATGTAAGAAAAAATCTCGACCAATTTCTTATGTAAGTGAAGATAAAAAA 232
Db 3899 ATATAAAATTTATGCCAACCCGCACI-----CAATCTTTCTGAAGAAGTTGATAAGTTAA 3952
Qy 233 ACAGTGTGTTGCGATAGCTCATTAGATATCGCAACACAGCTCTTACTTTTCCACGTAATAA 292
Db 3953 ATCGTAAAGCTCTTATTGATATTGATGACAGCAACAAAAATAATGTCTTAACGCCAAG 4012
Qy 293 AATTAACTCTATCCGAATCAGAAAAAAGCAAGTTTAAAAAATAAATGGAATATCTATCACTG 352
Db 4013 CAATTAGTCTGAACGAGGTTGAAAAAATAAATAAAGCAAAATACCGAGAAATAACCG 4072
Qy 353 GGAATAAATCGAGAACGCGAGAAATCAGAAAGTGAAGCTAGTACCCAAAGATTTTCCCTA 412
Db 4073 CAAGAATAATCAGACGGGGGAGTTAAAGAAAGTGAACCCATTCCTTTAGATTGGCCTA 4132
Qy 413 AAGATCTTTGTTCTGCTCCATTGCCAGATCATGTTAATGATTTTACATGTTGATACAAATC 472
Db 4133 GTGATTTAACTTTACCGCGTTACCTGAGAGCACAAACGATTATGTTTGGCGGGGAAA 4192
Qy 473 GAAAAA-----AAGCTTAGGTATAAAGCCTGTAAATAAAGATATATCGGCTCTTCTATTAT 527
Db 4193 GAAAAAGAGCTTGATGATTATCCAAGAAAAACAGTTAAT--CAATTGACGGCTTAGTATTGT 4250
Qy 528 TATTCCTACATTTAATCGTAGCGGTATTTTAGATATAACGTTAGCTGTTTGGTCAATCA 587
Db 4251 AATTCTACATATATATCGAGCAAAAATATCTTCAATTTACACTGCTGCTTTGTAAACCA 4310
Qy 588 GAAAAAATACTACCCATTTGAAAGTCTGTTGTCAGATGATGTTAGTAAAGAAAAAATTTACT 647
Db 4311 AAAGACCATATACGACTATGAAGTTATTCTTCCGATGATGGAAGTAAAGAAAAATATGA 4370
Qy 648 TACCATTGTGCAAAAATACGAAACAAACTTGACATAAAAGTATGTAAAGACAAAAAGATTA 707
Db 4371 AGAAATAGTAAGAGAAATTTGAAAGTTTATTAAATATAAAATATGTACGTCAGAGGATTA 4430
Qy 708 TGGATATCAATTTGTGTCAGTCAGAAACCTTAGCTTTACGTACAGAAAGTATGATCTTTGT 767
Db 4431 TGGATATCAACTGTGTGCTGTGTAGAAATCTTGGGCTTAGGCTGCAAGAGTAAATATGT 4490
Qy 768 CTCGATTTCTAGACTGCGATATGCGACCAACAAATTTATGGTTTCAATCTTATCTTACAGA 827
Db 4491 TCGAATTTCTGGATTTGTGATATGCTCCGAAACCCACTATGGTTTCACTATATATGAAC 4550
Qy 828 ACTATTAGAAGACAAATGATATTGTTTAAATTGGACCTAGAAAAATATGTGGTACTCATAA 887
Db 4551 ATTAGCGGTGGACGATAATGTTGCTCTAATTTGGCCCTAGAAAAATATATAGATACAAAG 4610
Qy 888 TATTACCGCAGAACAAATTCCTTAACGATCCATTTAATTAAGATATCACTACCTCAAAACCG 947
Db 4611 GCATACATATTAGATTTCCTTTCCCAAAATCACTAATAATGAATTCCTCGAAATCAT 4670
Qy 948 TACAATAACAATCTCTCGATTACATCAAAAGGAAATATATCGTTGGATTGGAGATTAGA 1007
Db 4671 TACTAATAATCAGGTTGCGAGCAAGGTTGAGCAAAACAAATCAGTTGACTGGCGAATAGA 4730
Qy 1008 ACATTTCAAAAAAACCGGATAATCTACGCTTATGTGATTTCTCGGTTTCTGTTTCTTTCG 1067

Db 4731 ACATTTCAAAAATACCGATAATCTAAGATTATGCAACACACCATTCGATTTTTCAGCGG 4790

Qy 1068 GGGTAATGTTGCATTTTCTAAAGAATGGCTAAATAAAGTAGGTGGTTCGATGAAGAATT 1127

Db 4791 AGGTAATGTCGTTTTCGGAAAAATGGCTTTTCCGTGCAGGATGGTTTGTGATGAAGATT 4850

Qy 1128 TAATCAATGGGGGGGGAAGATGTAGAATTTGGTTACAGATATTTGGTCCCAAGGCTGTTT 1187

Db 4851 TAGCGAATGGGGGGGAGGATAATGAGTTGGATATCTCTCTACAGAGAGGATGTTA 4910

Qy 1188 TTTTCAGAGTAATTTGACGGCGGAATGCCATCCATCAAGAACACCTGGTAAAGAAAATGA 1247

Db 4911 CTTTCGGTCTGTGAAGGAGCATGCGCATATCATCAAGAACACCCCGGGAAGAAAACGA 4970

Qy 1248 AACAGAACCGAAGCTGGTAAAGTAGTATACGGCTTAAATTTGTGAAAGAAAGTACCTTA 1307

Db 4971 GACGGATCGTGGCGGAGGAAAAATATTACTGTCTAAATTTGTTACAGCAAAAAGTCTCTTA 5030

Qy 1308 CATCTATAGAAAGCTTTTACCAATAGAAGATTCCACATATTCATAGATATCCTTTAGTTTC 1367

Db 5031 TTTCTATAGAAAAAAGAAAAATAGAACCGGCACATTAANAAGAGTACCACCTAGTATC 5090

Qy 1368 TATTTATATCCCGCTTATAACTGTGCAATATATTTCAAGATGTGTAGATAGTGTCT 1427

Db 5091 TATATATATCCCGCTTATAACTGTCTAAATATATTGTGTTGTGTTGAAAGCGCCCT 5150

Qy 1428 TAATCAAACTGTTGTCGATCTCGAGTTCGTATTGTAACGATGGTTTCAACAGATATATC 1487

Db 5151 TAATCAGACAAATAACTGACTTAGAGATGATGCATATGCGATGATGGTTTCCACAGATATC 5210

Qy 1488 CTTTGAAGTGTATCAATAGCTTTTGTGTAATAATCTTCCATAGGTCAGGCATCATGTAAACC 1547

Db 5211 ATTCCGATTTCTTCAGGAGCATATGCAACACCATCTCGAGTTGTTTATTTTACAAAA 5270

Qy 1548 AAATGGCGGAATAGCCTCAGATCAAAATGACGCGTTTCTTTTCTTAAAGTTATTACAT 1607

Db 5271 AAACAAAGGAATGGTTGACGATCTAATACAGCATAGTAGTTGTGTCGGGATTCATAT 5330

Qy 1608 TGGCAGCTATGATTCAGATGATTTATCTGAGCCTGATGCGATTTGAACTGTCTTTAAAGA 1667

Db 5331 AGGTCAGTTAGACTCTGATGACTTCTTGAACAGATGCTGTTGCACTATGCTAGATGA 5390

Qy 1668 ATTTTAAAGATAAAGCTAGCTTGTGTTTATACCACTAATAGAACGTCATTCGGA 1727

Db 5391 ATTTAGAAAGATCTATCATTTGGCATGTGTTTATACAACTAACCGTATATAGATCGTGA 5450

Qy 1728 TGGTAGCTTATTCGTTATGTTTACATTTGCGCAGATTTTTCACGAGAAAACTCACAC 1787

Db 5451 AGGTAATTTGATATCAAAATGGCTATAATGGCCCATTTTTCGGGAGAAAACTTACTAG 5510

Qy 1788 GGCTATGATTCCTCACCATTTTAGAATGTTTACGATTTAGAGCTTTGGCATTTAACGGATGG 1847

Db 5511 TGCAATGATATCATCTATTCAGGATGTTTCAAGAGAGATGGAACCTTAACCTGAAGG 5570

Qy 1848 ATTTAACGAAATATTGAACGCGGTGGATATGACATGTTTCTTAACTCAGTGAAGT 1907

Db 5571 TTTCAACGAATCGATCAGCAACGCGAGTTGATTACGATATGTTTAAACCTTAGTGAAGT 5630

Qy 1908 TGGAAAAATTAACATCTTATAAAATCTGCTATAACCGCGTATTACATGTTGTGATGAACAC 1967

Db 5631 TGGACCGTTCAAGCATATAACAAATTTGTTATTAATCGCGTATTGCAATGGTGAATAATC 5690

Qy 1968 ATCCATTAAAGAACTCGGCATTTCAAAAGAAAAACCATTTTGTGTAGTCAATCAGTCAAT 2027

Db 5691 GTCTATAAAAGTTGGATATTTCAAAAGGAAAAATCATTTTAAAGTTGTTAAAGAAATCAT 5750

Qy 2028 AAATAGACAGGCATCAATATTAATTAATGACAAATTTGATGATTTAGATGAAGATAG 2087

Db 5751 AAGTAGGCTAGGCATAAAAAATAAATAATTTTCAACATTAACTAATTTGAATGAATGTAG 5810

Qy 2088 AAGTATATCTTCAATAAACCGCTGAATATCAAGAGAAAAATG 2130

Db 5811 AAAATATCTCGGAAAAAATAGAGNAATGATTTATAATTAATTG 5853

RESULT 7

US-10-216-289-1

; Sequence 1, Application US/10216289

; GENERAL INFORMATION:

; APPLICANT: NINOMIYA, TOSHIO

; APPLICANT: SUGIYURA, NOBUO

; APPLICANT: KIMATA, KOJI

; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME

; FILE REFERENCE: 226882US0

; CURRENT APPLICATION NUMBER: US/10/216,289

; CURRENT FILING DATE: 2002-11-04

; PRIOR APPLICATION NUMBER: JP 2001-244685

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: JP 2001-324127

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: JP 2002-103136

; PRIOR FILING DATE: 2002-04-05

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 2058

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2058)

; OTHER INFORMATION:

US-10-216-289-1

Query Match 27.1%; Score 807.2; DB 6; Length 2058;

Best Local Similarity 63.2%; Pred. No. 3.le-145; Indels 13; Gaps 3;

Matches 1296; Conservative 0; Mismatches 743;

Qy 61 ATGATACATTTATCACAGCAATAAAGCATATACAGCAATGACTATGAATTAGCACTC 120

Db 1 ATGAGTATCTTAATCAAGCAATAAATTTATATAAAACAAAAATTCGCGCAAGCTTTA 60

Qy 121 AAATTTATTTGAGAAGTCTGCTGAACCTTACGGGCGAAAAATCGTTGAATTCCAAATTC 180

Db 61 TCTCTTTTGAAGAAGTTGCTGAAATTTATGATGTTAGTTGGGTGCGAAGCAAAATATAAA 120

Qy 181 AAATGTAAGAAATAATCTCGACCAATCTTATGTAAGTGAAGATAAAAAACAGTGT 240

Db 121 TTATGCCAACCGGCACT-----CAATCTTCTGAAGAAGTTGATAAGTTAATCGTAA 174

Qy 241 TCGATAGCTCATTTAGATATCGCAACACAGCTCTTTACTTTTCCACGTAACCAATTAAC 300

Db 175 GCTGTTATTGATATTGATGTCAGCAACAAAAATAATGTTCTTACGCCCAAGCAATTAG 234

Qy 301 CTATCCGAATCAGAAAAAACAAGTTTAAAAATAAATGAAATCTATCACTGGGAAAAAA 360

Db 235 CTGAACGAGGTTGAAAAAATAAATAAGCAATAATACCGAATAACCGCAAGAA 294

Qy 361 TCGGAGAACGCAAGAAATCAGAAAGGTGGAACTAGTACCCAAAGATTTTCTTAAAGATCT 420

Db 295 TCAGAACGGCGGAGTTAAAGGAGTCGAACCCATTCCTTTAGATTGGCTTAGTATT 354

Qy 421 GTTCTGTCTCAATGCGCAGATCATGTTAATGATTTTACATGTTGATACAAAATCAGAAA 479

Db 355 ACTTTACCGCGCTTACCTGAGAGCAACAAACGATTTATGTTGGGCGGGGAAAGAG 414

Qy 480 ----RAGCTTAGGTATAAGCCTGTAATAAAGAAATCGGTCTTTCTATTATTATTCCTA 535

Db 415 CTGTGATGATTTCCAGAAAAACAGTTAAT--CAATCAGCGGCTTAGTATTGTAATTCCTA 472

Qy 536 CATTTCAATCGTAGCCGCTATTTTAGATATAACGTTAGCCTGTTTGGTCAATCAGAAAA 595

Db 473 CATATAATCGCAAAAAATACCTTGCAATTCAGCTTCTTCTTCTTCTTCTTACCAAAAGACCA 532

Qy 596 ACTACCACTTTGAAGTCGTTGTTGAGATGATGGTAGTAGGAAAACTTACTTACCATTG 655

Db 533 TATACGACTATGAAGTTTATTGTTGCCGATGATGGAAGTAAGAAAAATATTGAAGAAATAG 592
Qy 656 TGC AAAAATACGACAAAAAATCTTACATTAAGTATGTAAGCAAAAAGATTTATGATATC 715
Db 593 TAGAGAAATTTGAAGTTTATAAATATAAATATGTAAGTACGTCAGAAGATTTATGGATATC 652
Qy 716 AATTGTGTGAGTACGAAACCTTAGTTTAGTACAGCAAAAGTATGATTTGTCTFCGATTC 775
Db 653 AACTGTGTGCTGTAGAAATCTTGGGCTTAGGGCTGCAAGATATAATATGTTCGAATTC 712
Qy 776 TAGACTCGGATATGGACACACAAATATATGGGTTCAATCTTATCTTACAGAACTATTAG 835
Db 713 TGGATTTGATATGGCTCCGAACCACTATGTTGTTCACTATATATGGAACATTTAGCGG 772
Qy 836 AAGCAATGATATTTGTTTAAATGGACCTAGAAAAATATGTTGATACATCAATAATATTACG 895
Db 773 TGGAGTAATGTTGCTCTAATGGCCCTAGAAAAATATATAGATACAAAGCAGCATACAT 832
Qy 896 CAGAACAATTCCTTAAGCATCCATATTTAATAGAAATCACTACCTGAAACCGCTACAAATA 955
Db 833 ATTTAGATTTCTTCCCAAAAATCACTAATAATGAATTCCTGAAATCAATTACTAATA 892
Qy 956 ACAATCCTTCGATTAACAAAAGAAATATATCGTTGGATGAGATGAGAACTATTCATA 1015
Db 893 ATCAGGTTGCGAGCAAGGTTGAGCAAAACAAATCAGTTGACTGGCGAATAGAACATTTCA 952
Qy 1016 AAAAAACCGATATACGTCTATGATCTCCGTTTCTGTTATTTTGTTCGGGTTAATG 1075
Db 953 AAAATACCGATAATCTAAGATATTGCAACACACACCATTTTCGATTTTTTAGCGGAGGTAA 1012
Qy 1076 TTGCATTTTCTAAGAAATGGCTAAATAAAGTAGTTGGTTTCGATGAAGAAATTTAATCAT 1135
Db 1013 TCAGTTTTCGAAAAAATGCTTTTCGTCGAGGATGGTTTGATGAGAGTTTACGCAT 1072
Qy 1136 GGGGGCGGAGATGTAAGATTTGGTTTACAGATTTATTTGCCAAAGGCTGTTTTTCAGAG 1195
Db 1073 GGGGGGGGAGGATAATGAGTTTGGATATGCTCTACAGAGAAGATGTTACTTTTCG 1132
Qy 1196 TAATTGACGGGGAATGCCATCATCAAGAACCCATCGTTGAAGAAATGAACAGAAC 1255
Db 1133 CTGTTGAAGAGCAATGGCATATCATCAAGAACCCCGGGAAGAAACGAGACGGATC 1192
Qy 1256 GCGAAGCTGTGTAAGATTTACGCTTAAAAATTTGTGAAGAAAGGTACCTTACATCTATA 1315
Db 1193 GTGCGGAGGAAAAATATTACTGTTCAATTTGTTACAGCAAAAGTTCTTATTTCTATA 1252
Qy 1316 GAAAGCTTTTACCAATAGAAATTCATATTCATAGAAATACCTTTAGTTTCTATTATA 1375
Db 1253 GAAAAAAGAAAAATAGAATCCGCGACATTTAAAAAGAGTACCCACTAGTATCTATATA 1312
Qy 1376 TCCCGCTTATACTGTGCAAAATTTATATCAAGATGTGATAGTGTCTCTTAATCAAA 1435
Db 1313 TCCCGCTTATACTGTCTAATATATATGTTCTGTTGTGAAGCGCCCTTAATCAGA 1372
Qy 1436 CTGTTGCTGCTCGAGGTTTGTATTTGTAACGATGTTCAACAGATAATACCTTAGAAG 1495
Db 1373 CAATAACTGACTTAGAATATGATATGCGATGATGTTCCACAGATGATACATTCGCGA 1432
Qy 1496 TGATCAATAGACTTTTATGTAATTAATCCTAGGGTACGATCATGCTTAAACCAATGGCG 1555
Db 1433 TTCTTCAGAGACATTTATGCAACCATCTCCGAGTTTCGTTTTTATTTTCAAAAAAACAAG 1492
Qy 1556 GAATACCTCAGCATCAATGCAAGCCCTTCTTTTGTGTAAGGTTTATACATTTGGCGAGT 1615
Db 1493 GAATTTGTTTACGATCTAATACAGCATGTAGATTGTGTCGGGGGATTTCTATATAGTCAGT 1552
Qy 1616 TAGATTTCAGATGATTATCTTAGGCTGATGCGAGTTGAACTGTGTTTAAAAAGAAATTTTAA 1675
Db 1553 TAGACTCTGATGACTTTCTTGAACACAGATGCTGTGTAACATGCTCTAGATGAATTTAGAA 1612
Qy 1676 AAGATAAAACGCTAGCTGTTGTTTATACCACTAATAGAAACGTCATCCCGGATCGTAGCT 1735
Db 1613 AAGATCTATCATTTGGCATGTTGTTTATACAACTAACCCGCTAATATAGATCGTGAAGGTAAT 1672

Qy 1736 TAATCGCTTAATGGTTACAAATTTGGCAGAAATTTTTCACGAGAAAAAATCTCACACGGCTATGA 1795
Db 1673 TGATATCAAAATGGCTATTAATTTGGCCCATTTTATTCGGAGAAAAAATCTTACTAGTCAATGA 1732
Qy 1796 TTGCTCACCAATTTTAGAATGTTTACGATTTAGAGCTTTGGCATTTTAACGGATGGATTTAACG 1855
Db 1733 TATGTCATCAATTTTCAGGATGTTTCACAGCAAGAGCATGGAACCTTAAGGTTTCAACG 1792
Qy 1856 AAAATATTGAACGCCGCTGGATTTAGCATTTTCCCTTTAAACTCAGTCAAGTTGGAATAAT 1915
Db 1793 AATCGATCAGCAAGCAGCTTTGATTTACGATATGTTTAAAACTTTAGTGAAGTTGGCCGT 1852
Qy 1916 TTAACATCTTAAATAAATCTGTATACCGGTATTACATGGTGTATTAACACATCCATTATA 1975
Db 1853 TCAAGCATATAACAAATTTGTTATATCCGTTATTCGATGGTGGGAAATAGCTCTATAA 1912
Qy 1976 AGAACTCGGCATTCAAAGAAAAACCAATTTTGTGTAGTCAATCAGTCAATTAATAGAC 2035
Db 1913 AAAAGTTGGATATTCAAAAGGAAAAATCATTTTAAAGTTGTTAAACCAATCATTTAAGTAGGC 1972
Qy 2036 AAGCATCAATTTATTAATTTATGACAAATTTTGATGATTTTAGATGAAAGTAGAAGATATA 2095
Db 1973 TAGGCATAAAAAAATATAAATATTACCATTAACTAATTTGAATGAATGTAGAAATATA 2032
Qy 2096 TCTTCAATAAAA 2107
Db 2033 CTGGGAAAAAAA 2044

RESULT 8
US-09-134-000C-2987
; Sequence 2987, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2987
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2987

Query Match 2.9%; Score 86; DB 5; Length 993;
Best Local Similarity 54.0%; Pred. No. 8e-08;
Matches 176; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy 1354 ATACCTTTAGTTTCTTATTTATATATCCCGCTTATATCTGCAAAATTTATTTCAAAAGATGT 1413
Db 19 ATGCCCAAAATTAGTATTATTGTTCTCTATACATGTAGAAAAATATTTAGAAAAATGT 78
Qy 1414 GTAGATAGTCTTTAATCAAACTGTTTCTGATCTCGAGGTTTGTATTGTAACGATGT 1473
Db 79 GTACGCTCTATTTTAGCTCAAAACGTTTACTGACTTTGAATTAATTTCTGCTGACGATGC 138
Qy 1474 TCAACAGATAATACCTTAGAAGTGTATCAATAGCTTTTATGTTAATTAATCTAGGTCAGC 1533
Db 139 TCTCCAGACAGTCTTGGAGCAATGTGTGATCAGTTTGTGAAACAGATCAACGGGTAAAA 198
Qy 1534 ATCATGCTTAACCAAAATAGCGGAATAGCTCAGCATCAAAATGACGCGCTTTCTTTTGTCT 1593
Db 199 GTTATCCATAAAGAAATAGTGGGCTAAGCGATGCTGTAATGCTGGAAATGAATAGCA 258
Qy 1594 AAAGTTTATTAATTTGGCAGTTAGATTTAGATGATGATTATCTTTAGCCCTGATGACGTTGAA 1653
Db 259 ACAGGTGAGTATTTAGGTTTCTGTAGATAGTATGATGATTACATTTGCAGATGATATGTATGAA 318

QY 1654 CTGCTGTTAAAGAAATTTTAAAGA 1679
|| | | | | | | | | | | | | | | | |
Db 319 CTATTATACAAATATCGTAAAGA 344

RESULT 9

US-10-303-161-1/c
; Sequence 1, Application US/10303161
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-161-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 1296 AAGGTACCTTACATCTATAGAAGCTTTTACCAATAGAAGATTCACATATTCATAGAAT 1355
|| | | | | | | | | | | | | | | | |
Db 11424 AATTTCTTAACCTTGTATAGAATGTTTATATTTTAAACAAATTTAGGAAATAT 11365

QY 1356 ACCTTTAGTTCTTATATATCCCGCTTATACTGTGCAAAATATATCAAGATGTGT 1415
|| | | | | | | | | | | | | | | | |
Db 11364 GCCACAACCTTCTATCATATATCCCGCTTTTAAATCTTGGCATTTATCTCAAGAGCTTT 11305

QY 1416 AGATAGTCTCTTAATCAAACTGTTCGATCTCGAGTTTGTATTTGTAACTGATGTTTC 1475
|| | | | | | | | | | | | | | | | |
Db 11304 GCAAGTGTATAAATCAGACTTTTAAAGATATTGAAATTTGATTCGATGATAAAG 11245

QY 1476 AACAGATAAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATAATCCTAGGGTACGCAT 1535
|| | | | | | | | | | | | | | | | |
Db 11244 TAAAGATAATAGTTTAAATAGTTTGTAGATTTTGCACCAAGATATAAAT 11185

QY 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATCGACCGTTCTTTTTCG 1592
|| | | | | | | | | | | | | | | | |
Db 11184 CTTTCAAAATGAAGAAATTTAGGCACCTTTTGCAGTAGAATTTGGGGGTTTTCGATTC 11125

QY 1593 TAAAGTTATTACATTTGGGCAGTTAGATTTCAGATGATTATCTTGAGCCTGATGAGTTGA 1652
|| | | | | | | | | | | | | | | | |
Db 11124 TAGTCTGATTTTAAATGTTTGTAGATGATGATTTTAAACACCTGATGCTTGGA 11065

QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTGTGTTTATACC 1704
|| | | | | | | | | | | | | | | | |
Db 11064 AATAGCGTTTAAAGAGATGAAAAAGGTTTGAATTTGCTTTGTTGTATGCC 11013

RESULT 10

US-10-303-162-1/c
; Sequence 1, Application US/10303162
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-162-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 1296 AAGGTACCTTACATCTATAGAAGCTTTTACCAATAGAAGATTCACATATTCATAGAAT 1355
|| | | | | | | | | | | | | | | | |
Db 11424 AATTTCTTAACCTTGTATAGAATGTTTATATTTTAAACAAATTTAGGAAATAT 11365

QY 1356 ACCTTTAGTTCTTATATATCCCGCTTATACTGTGCAAAATATATCAAGATGTGT 1415
|| | | | | | | | | | | | | | | | |
Db 11364 GCCACAACCTTCTATCATATATCCCGCTTTTAAATCTTGGCATTTATCTCAAGAGCTTT 11305

QY 1416 AGATAGTCTCTTAATCAAACTGTTCGATCTCGAGTTTGTATTTGTAACTGATGTTTC 1475
|| | | | | | | | | | | | | | | | |
Db 11304 GCAAGTGTATAAATCAGACTTTTAAAGATATTGAAATTTGATTCGATGATAAAG 11245

QY 1476 AACAGATAAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATAATCCTAGGGTACGCAT 1535
|| | | | | | | | | | | | | | | | |
Db 11244 TAAAGATAATAGTTTAAATAGTTTGTAGATTTTGCACCAAGATATAAAT 11185

QY 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATCGACCGTTCTTTTTCG 1592
|| | | | | | | | | | | | | | | | |
Db 11184 CTTTCAAAATGAAGAAATTTAGGCACCTTTTGCAGTAGAATTTGGGGGTTTTCGATTC 11125

QY 1593 TAAAGTTATTACATTTGGGCAGTTAGATTTCAGATGATTATCTTGAGCCTGATGAGTTGA 1652
|| | | | | | | | | | | | | | | | |
Db 11124 TAGTCTGATTTTAAATGTTTGTAGATGATGATTTTAAACACCTGATGCTTGGA 11065

QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACCGCTAGCTGTGTTTATACC 1704
|| | | | | | | | | | | | | | | | |
Db 11064 AATAGCGTTTAAAGAGATGAAAAAGGTTTGAATTTGCTTTGTTGTATGCC 11013

RESULT 11

US-10-303-118-1/c
; Sequence 1, Application US/10303118
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474

; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-118-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;
Qy 1296 AAAGGTACCTTACATCTATAGAAAGCTTTTACCATAAGAGATTCACATATTCATAGAAAT 1355
Db 11424 AAATTTCTTAAACCTTGATAGAAATTTTATTATTTTAAACAAAATTTAGGAAAATAT 11365
Qy 1356 ACCTTTAGTTCTTATATATCCCGCTTATAAAGTGTGCAAAATATATTCAAAGATGTG 1415
Db 11364 GCCACAACCTTCTATCATATCCCGCTTTTAAATCTTTGCGATTTTATCTCAAGAGCTTT 11305
Qy 1416 AGATAGTCTCTTAACTCAACCTGTCGATCTCAGGTGCTATTTGTAATTCATGATGTTTC 1475
Db 11304 GCAAGTGTGTAATACAGCTTTTAAAGATATTTGAAATTTGATATCGATGATAAAG 11245
Qy 1476 AACAGATAATACCTTAGAAGTGTGTCGATCTCAGGTGCTATTTGTAATTCATGATGTTTC 1535
Db 11184 TAAAGATAATAGTTTAAATATGTTTATGAAATTTGCAAAAAGATCCCAAGATAAAT 11185
Qy 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATGCAGCGCTTTCTTTTGC 1592
Db 11184 CTTTCAAAATGAAGAAAATTTAGGCACCTTTGCAAGTAGAAAATTTGGGGGTTTTCATTC 11125
Qy 1593 TAAAGGTTTATACATGCGCAGTTAGATTCAGATGATTTCTTGAGCCTGATGCAATTC 1652
Db 11124 TAGTCTGATTTTAAATATGTTTATGATAGTAGTATGATTTTAAACACCTGATGTTGCA 11065
Qy 1653 ACTGTGTTTAAAGAAATTTTAAAGATGATAAAGCTAGCTGTGTTTATAC 1704
Db 11064 AATAGCGTTTAAAGAGATGAAAAAGGTTTTCATTTGCTTTGTTGATGCC 11013

RESULT 12

US-10-303-128-1/c
; Sequence 1, Application US/10303128
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus

US-10-303-128-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;
Qy 1296 AAAGGTACCTTACATCTATAGAAAGCTTTTACCATAAGAGATTCACATATTCATAGAAAT 1355
Db 11424 AAATTTCTTAAACCTTGATAGAAATTTTATTATTTTAAACAAAATTTAGGAAAATAT 11365
Qy 1356 ACCTTTAGTTCTTATATATCCCGCTTATAAAGTGTGCAAAATATATTCAAAGATGTG 1415
Db 11364 GCCACAACCTTCTATCATATCCCGCTTTTAAATCTTTGCGATTTTATCTCAAGAGCTTT 11305
Qy 1416 AGATAGTCTCTTAACTCAACCTGTCGATCTCAGGTGCTATTTGTAATTCATGATGTTTC 1475
Db 11304 GCAAGTGTGTAATACAGCTTTTAAAGATATTTGAAATTTGATATCGATGATAAAG 11245
Qy 1476 AACAGATAATACCTTAGAAGTGTGTCGATCTCAGGTGCTATTTGTAATTCATGATGTTTC 1535
Db 11244 TAAAGATAATAGTTTAAATATGTTTATGAAATTTGCAAAAAGATCCCAAGATAAAT 11185
Qy 1536 CAT---GTCTAAACCAATGCGGAATAGCCTCAGCATCAATGCAGCGCTTTCTTTTGC 1592
Db 11184 CTTTCAAAATGAAGAAAATTTAGGCACCTTTGCAAGTAGAAAATTTGGGGGTTTTCATTC 11125
Qy 1593 TAAAGGTTTATACATGCGCAGTTAGATTCAGATGATTTCTTGAGCCTGATGCAATTC 1652
Db 11124 TAGTCTGATTTTAAATATGTTTATGATAGTAGTATGATTTTAAACACCTGATGTTGCA 11065
Qy 1653 ACTGTGTTTAAAGAAATTTTAAAGATGATAAAGCTAGCTGTGTTTATAC 1704
Db 11064 AATAGCGTTTAAAGAGATGAAAAAGGTTTTCATTTGCTTTGTTGATGCC 11013

RESULT 13

US-10-303-134-1/c
; Sequence 1, Application US/10303134
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-134-1

Query Match 2.7%; Score 79.2; DB 6; Length 11474;
Best Local Similarity 51.2%; Pred. No. 1.8e-06;
Matches 211; Conservative 0; Mismatches 198; Indels 3; Gaps 1;
Qy 1296 AAAGGTACCTTACATCTATAGAAAGCTTTTACCATAAGAGATTCACATATTCATAGAAAT 1355
Db 11424 AAATTTCTTAAACCTTGATAGAAATTTTATTATTTTAAACAAAATTTAGGAAAATAT 11365
Qy 1356 ACCTTTAGTTCTTATATATCCCGCTTTATAAAGTGTGCAAAATATATTCAAAGATGTG 1415

Db 11364 GCACAACTTCTATCATATCCCGCTTTTAACTTCGCGATTATATCTCAAGAGCTTT 11305
QY 1416 AGATAGTGTCTTAATCAAACTGTGTGTCATCTCGAGGTTTGTATTTGTAAAGCATGGTTC 1475
Db 11304 GCAAACTTGTATTAATACAGATTTAAAGATATTTGAAATTTTGAATATCGATGATAAAAG 11245
QY 1476 ACAGATATACCTTGAAGTGTATCAATAAGCTTTATGCTAATAATCCCTAGGTAGCGCAT 1535
Db 11244 TAAAGATAATAGTTTAAATATGTTTGAATTTGCAAAAAGATCCCAAGATAAAAT 11185
QY 1536 CAT--GTCTAAACCAATGGGGAGTAGCCTCAGCATCAAAATGACCGCTTTCTTTTGC 1592
Db 11184 CTTTCAAAATGAAGAAATTTAGGCACCTTTCGAAGTAGAAATTTGGGGGTTTTCGATTC 11125
QY 1593 TAAAGTATTATACATGGGAGTTAGATTCAGATGATTTATCTTGAGCCCTGATGCACTTGA 1652
Db 11124 TAGTCTGATTTTAAATGTTTGTAGATGATGATTTTAAACACCTGATGCTGTGA 11065
QY 1653 ACTGTGTTTAAAGAAATTTTAAAGATAAAACGCTAGCTTGTGTTTATACC 1704
Db 11064 AATAGCGTTTAAAGAGATGAAGAAAGGTTTGTGTTTGTGTTTGTATGCC 11013

RESULT 14
US-10-240-454-27/c
; Sequence 27, Application US/10240454
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Metabolism
; FILE REFERENCE: 5013.1010
; CURRENT APPLICATION NUMBER: US/10/240,454
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/04016
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 68
; SEQ ID NO 27
; LENGTH: 17527
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-454-27

Query Match 2.6%; Score 77.2; DB 6; Length 17527;
Best Local Similarity 43.9%; Pred. No. 4.3e-06;
Matches 434; Conservative 0; Mismatches 543; Indels 11; Gaps 2;
QY 1992 AAAGAAACACATTTGTTAGTCAATCAGTCATTAATAGACAAAGGCATCAATTAATA 2051
Db 16996 ATAAAAAACCTGATCTCTTAATCTTAAACATTTAAAAAATACTAAACACACGA 16937
QY 2052 TAATTATGACAAATTTGATGATTTAGATGAAAGTAGATATCTTCAATAAACCGC 2111
Db 16936 TAATCAACCTAAATCCCACTACTCAAAAAACAAAAATAAATAATATTACTTAAACCC 16877
QY 2112 TGAATATCAAGAAATGGATATGTTAAAGATCTTAAACCTCAATCAAAATAAAGATGC 2171
Db 16876 AAAATTTAAATTAACAATAAACCAAAATCACACCCTATATCTTCAACCTAAACACAAA 16817
QY 2172 CAAATCGCAGTCAGTATTTCTTATCCCAATACATTAACGCGTTAGTGAATAAACTAAA 2231
Db 16816 CAAAACCCCACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16757

QY 2232 CAATATTATGAATATAATAAAATATATTCTGTTATTTATTTCTACATGTTGATAAGATCA 2291
Db 16756 AAAAAACAAACAAACAAAAAACTTCCCAATAAT-----CACTATTAAACCTAAAA 16701
QY 2292 TCTTACACAGACATCAAAAAAGAAATATTGGCTTTCTATCATAGCACCAGTGAATAT 2351
Db 16700 AAAAAAATAAAAAAACCACAAACAAACACCAACCAATACACCTCTTAAAC 16641
QY 2352 TTTTACTAAATATGACATCTCATATTACAGAGTAAATAGACATAAATAAACTGAGGCACA 2411
Db 16640 TTTTCAAAATATAGGACTCAATAAAAAACAAATAATTTCTATAACAAC-----AT 16588
QY 2412 TTTAAGTAAATATAAATTAAGTCAGTTAAATCTAAATTTGTAATACATCATTTTGA 2471
Db 16587 TTTTAAATAAATTAATAATTTTATACACTATTTTAAATAATCGAACACCACTAACTCA 16528
QY 2472 TAATCATGACAGCTATTCTGTTAAAAATGACAGCTATGCTTATATGAAAAATATGATGT 2531
Db 16527 TACCTATAATCTTAAACCTTTTAAATAATCAACAATAATCACCCTAAATCAAAAT 16468
QY 2532 CGGCATGAATTTCTCAGCATTAACACATGATGATGATCGAGAAATCAATCGGCATCCACC 2591
Db 16467 TAGCACCAACTAACCAAAATAACGAAATCCCGTCTCTACTAAAAACACACAAAAAA 16408
QY 2592 ATTTAAAAAGCTGATTAACACCTATTTTATGACATGACTTAAGAGTATGATGTGAA 2651
Db 16407 TTAACGAAACATATAACACACACCTATAATCCCACTACTCAAAAAACCTAAATAAAT 16348
QY 2652 AGGGCATCAAGGTATGTTTATGAGTATGGCTACCGCATGAGCTTCTGACGATTTAT 2711
Db 16347 TACTTAAACCCAAAAATAAATAATTAATAACCTAAATCTTACCATTACACTCCAACC 16288
QY 2712 TAAAGAACTCATCACATCTCGCCAATCAATGATGATGTCGACAGATATACACTGAGGA 2771
Db 16287 TAAACAAACAAACGAAACCTCTCAAAAAAATAATTAATTTATTTAAATTTAAAAA 16228
QY 2772 TATTGGTTCCATTTGCTGACCTTTTAACTTAGAAAAAGAAACCGCCATGTTTAAATA 2831
Db 16227 ATATAATCTCTATATAAATAATAAATAATAAATAACAAAAATAATTTATAAATAA 16168
QY 2832 AACATCGACCTGACTTATATGCTTGGGACGAAATTTCAATGACAAATGACAAAT 2891
Db 16167 AACCTTTACCAATAAAAAACCAAAAAACAAAAAATAAATAAATAAATAAATAAATA 16108
QY 2892 TCAAGTGCAAAAAGCGGAAATATCCCGTTTAAACAGTTTCAATTTAATAGTATAAC 2951
Db 16107 TATCCCTATTAAAAATCAAAACTATCTCTCATATATTCCCTTCAACACATAAATAA 16048
QY 2952 GCTATAAACATTTGCAATTTTATTAAAA 2979
Db 16047 CAAAAATATTATCCCAATTTTATAACA 16020

RESULT 15
US-10-264-213-27/c
; Sequence 27, Application US/10264213
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Lubbers, Mark William
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043c3
; CURRENT APPLICATION NUMBER: US/10/264,213
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 13825
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus

US-10-264-213-27

Query Match	2.68;	Score 76.6;	DB 6;	Length 13825;
Best Local Similarity	53.58;	Pred. No. 5.6e-06;		
Matches 160;	Conservative 0;	Mismatches 139;	Indels 0;	Gaps 0;
Qy 1354	ATACCTTTAGTTCTTATATATATCCCGCTTTATAACTGTGCAAAATATATATCAAGATGT	1413		
Db 11874	ATGCCTTTAGTTCTTATAAATCGTTCCCTATTATATGTTGAAAAAATCTTACAGCGATGC	11815		
Qy 1414	GTAGTAGTGCTCTTAATCAACAGTGTCTCGATCTCGAGGTTTGTATTTGTAACGATGTT	1473		
Db 11814	ATAGACAGTTTATTTGGCCAAACATACGTCAAATATAGAGATAGTCCCTAGTTGACGACCGA	11755		
Qy 1474	TCAACAGATAATACCTTTAGAAAGTGATCAATAAGCTTTATGGTAATAATCTCTAGGATAGC	1533		
Db 11754	TCGCGACAGCCGTTCACTGTCTAATCTGTAAAAAAAATGCTCAGCAGAGTGAACGAGTGCGA	11695		
Qy 1534	ATCATGTCTAAACCAAAATGGCGGAATAGCCCTCAGCATCAAAATGCAGCGCTTCTTTTGCT	1593		
Db 11694	GTCTTCTCAAGACAAATGGTGGGGTTGCAGATAGCGGTATTTTGGTGTATTCGGTTGCA	11635		
Qy 1594	AAAGTTTATTACATTTGGGCACTTAGATTTCAGATGATTATCTTTGACCCCTGATGCAGTTGA	1652		
Db 11634	CGGGCAACAATGGATATCATTTGTGTGTCGCCCATGATATGTGGATCCAGATTATATTCGA	11576		

Search completed: January 4, 2003, 02:18:55
Job time : 319 secs

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 21:14:45 ; Search time 3682 Seconds

(without alignments)

13103.301 Million cell updates/sec

Title: US-09-842-484A-1

Perfect score: 2979

Sequence: 1 ttataaactgattaaagaag.....acatttgcatattataaaa 2979

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16134066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93.4	3.1	1101	17	CNS00EVL
c 2	86.8	2.9	910	13	BM415636
3	82.8	2.8	996	17	CNS00FHH
c 4	81.8	2.7	641	17	AQ946120
5	80	2.7	562	13	BI815821
6	78.8	2.6	1092	17	CNS020K7

c 7	77.8	2.6	1101	17	CNS00EVL	AL069706 Drosophil
c 8	77.6	2.6	1059	17	CNS002B	AL097133 Drosophil
c 9	77.4	2.6	1169	17	CNS06KHQ	AL402900 T3 end of
c 10	76.8	2.6	994	17	CNS04NOJ	AL298972 Tetraodon
c 11	76.4	2.6	973	12	BG850019	EG850019 1024027H0
c 12	76.2	2.6	778	17	AG058583	AG058583 Pan trogl
c 13	76	2.6	832	17	BH391984	BH391984 AG-ND-138
c 14	76	2.6	1101	17	CNS0021J	AL061936 Drosophil
c 15	75.8	2.5	1001	17	CNS0155H	AL105023 Drosophil
c 16	75.8	2.5	1101	17	CNS000B8	AL063632 Drosophil
c 17	75.6	2.5	829	17	CNS03LUB	AL249932 Tetraodon
c 18	75.2	2.5	1101	17	CNS0039G	AL063921 Drosophil
c 19	75.2	2.5	1101	17	CNS00E07	AL069440 Drosophil
c 20	75	2.5	668	9	AL514901	AL514901
c 21	75	2.5	1201	17	CNS0167M	AL106396 Drosophil
c 22	74.8	2.5	1139	17	AQ897537	AQ897537 HS_3153-A
c 23	74.6	2.5	894	17	CNS018BG	AL109126 Drosophil
c 24	74.2	2.5	1187	17	BI11102	BI11102 F19C22-T7 I
c 25	74	2.5	897	17	CNS07AB2	AL436389 T7 end of
c 26	74	2.5	945	17	CNS04DOK	AL285149 Tetraodon
c 27	73.6	2.5	781	17	CNS009D0	AL053444 Drosophil
c 28	73.2	2.5	869	17	AZ682338	AZ682338 ENTKR07TF
c 29	73	2.5	1135	17	CNS033GQ	AL226115 Tetraodon
c 30	72.8	2.4	942	17	CNS018GS	AL109318 Drosophil
c 31	72.8	2.4	1101	17	CNS0039G	AL063921 Drosophil
c 32	72.6	2.4	660	11	AY068550	AY068550 Schmidtea
c 33	72.4	2.4	1101	17	CNS00E07	AL069440 Drosophil
c 34	72.2	2.4	963	9	AL566565	AL566565
c 35	72	2.4	935	17	BI0881	BI0881 F24H6-Sp6.1
c 36	72	2.4	1027	17	CNS02T50	AL212733 Tetraodon
c 37	72	2.4	1055	14	BQ876453	BQ876453 AGENCOURT
c 38	71.8	2.4	1007	17	CNS06X9S	AL419462 T3 end of
c 39	71.6	2.4	665	12	BG854775	BG854775 1024040G1
c 40	71.6	2.4	884	12	BF272396	BF272396 GA_E8001
c 41	71.4	2.4	892	17	AZ539768	AZ539768 ENTDC17TR
c 42	71.2	2.4	1225	17	CNS0161D	AL106171 Drosophil
c 43	71	2.4	576	17	CNS035N7	AL228940 Tetraodon
c 44	71	2.4	928	17	CNS00DKY	AL071865 Drosophil
c 45	70.8	2.4	964	13	BM416130	BM416130 OP21217 M

ALIGNMENTS

RESULT 1	CNS00EVL	1101 bp	DNA	linear	GSS 04-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC:				
DEFINITION	BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL069706				
VERSION	AL069706.1	GI:4949849			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial				

EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .1101

/organism="Drosophila

/db_xref="taxon:7227"

```
/clone="BACR29B23"
```

```
/clone_lib="RPCI-98"
```

/note="end : T7"

BASE COUNT	419 a	91 c	60 g	299 t	232 others
BASE COUNT	419 a	91 c	60 g	299 t	232 others

BASE C
ORIGIN

Query Match 3.18; Score 93.4; DB 17; Length 1101;

Best Local Similarity 34.3%; Pred. NO. 4.7e-08;
Matches 226; Conservative 117; Mismatches 310; Indels 6; Gaps 1;

Q7 1062 AACACATCCATTAGAAACTCGGCATTCAAAAGAAAAACCATTTGTTGTAGTCAATCAG 2022

QY 1963 AACACATCCATTAAAGAACTCGGCATTCAAGAGAAAACCATTTGTTGTAATCATTCAG 2022

[illegible]

Db 449 HAMATYTCCTCAHTWTMMMMTMMWAAATWTWWAAAATAAATTAATWAAATWAAAAAANNNNNAI

0002 ECGECCACCCCECAATTATTTATGACAATAATTGATGATTAGATGAA 2082

QY 2023 TCATTAAATAGACAAGGCATCAATTATTAATATGACAAAATTGATGATTTAGATGAA 2082

[illegible]

Db 509 TTTTWWWWTTTWTTTTWAWWWTTTAAATTTAAAWWAAATAAATLAA 508

2142

QY 2083 AGTAGAAAGTATATCTTCAATAAAACCGCTGAATATCAAGAGAAAATGGATATGTTAAAA Z14

628 TATTTTAA

Db
Ddb

569 WAA TTTAWA AWWTATATA TTAA TWTA TAATAA TWWTATTAA TTTT TTTWATAA GCG

2202

QY 2143 GATCTTAACTCATTCATAAATAAGATGCCAAATTCGCAAGTCAGTATTTCTATCCCAAT 220

688

D_b 629 AATTTTAAATAATTATTAATAAATAAATAATTTTAAAAAT 688

2262

QY 2203 ACATTAAACGGCTTAGTGAAAAAACTAAACAATATTATTGAATATAATAAAAAATAATATTC 2262

742

Db 689 AAAAAAAAA-----AAAAAAAAATWAAWAAATWATAWATAAATWATAWATAAAAAAA / 42

2322

QY 2263 GTTATTCTACATGTTGATAAGAATCATCTTACACCAGACATCAAAAAAGAAATATTTG 2322

[illegible]

RESULT 2	EST 28-JAN-2002
BM415636/c	
LOCUS	910 bp mRNA linear
DEFINITION	Op20714 Mixed Stage EST's from <i>Globochorda pallida</i> , the potato cyst nematode <i>Globochorda pallida</i> cDNA. mRNA sequence.

ACCESSION	BM415636
VERSION	BM415636.1
KEYWORDS	EST.

D	b		659	GTTTAATMTAATACAWTAAATAWAAAAAAAATWWAAAAAAAATWAAAAAAAATWAAAAA	TAAA	718
Q	y		2378	ACAGGACTAATAGACTAATAAAAACTGAGGCACCATTTTTAAGTAAATATTAAATAAATTAAGTC	2437	
D	b		719	AAAAWAAATWAAAAAAAAATAWAAAAATAAWTAAATTWATMAAAAAAATAAAAAAAWATTT	778	
Q	y		2438	AGTTAAATCATAATTGGAATACATCAATTTTGGATAATCATGACAGCGCTATTCGTTAAAA	2497	
D	b		779	RWTATATATWAANWAAWTAAMWTTTKAWMATWAAATWAAAAAATAAAAAAAAHWATTAATAAA	838	
Q	y		2498	ATGACAGCTATGCTTATATGAATAAAAAATATGATGCGCATGAATTTCTCAGCATTAACAAC	2557	
D	b		839	AAAAAANAANAATWNTATWNTAAATYAAATTTAAATTTWAAAAAATWWWTTTTTTTTT	898	
Q	y		2558	ATGATTCGATCGAGAANAATCAATGCGCATCCACCATTATAAAAAAGCTGATTAACACCTATT	2617	
D	b		899	TTTTTTTTTYCCSSSCSSCSSSCSVAAAVAAATTAWAAATKTWATTTWATTTT	958	
Q	y		2618	TAAATGACAATGACTTTAAGAAGTAGTAA	2645	
D	b		959	KWAATWAAAAAATWATTTTTTTTATWAA	986	
RESULT 4						
AQ946120/c						
LOCUS						
DEFINITION						
Sheared DNA-46J23.TR Sheared bp DNA linear GSS 27-JAN-2000						
Sheared DNA-46J23, DNA sequence.						
ACCESSION						
AQ946120						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C., Gerard,C., Leech,V., de Jong,P., Ullu,E., Melville.S., Donelson.J., Fraser,C. and Adams,M. Determination of clone end sequences from Trypanosoma brucei GUTat 10.1 sheared DNA library Unpublished (1999) Other_GSs: Sheared DNA-46J23.TF Contact: Najib M. El-Sayed Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: nelsayed@tigr.org Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared DNA library constructed at TIGR. Clones will be available for distribution through ATCC. Sheared DNA end sequences search page: http://www.tigr.org/tdb/mdb/cbdb/ . Seq primer: M13-Reverse Class: shotgun.						
FEATURES						
source						
1..641						
/organism="Trypanosoma brucei"						
/strain="TREU927/4 GUTat 10.1"						
/db_xref="taxon:5691"						
/clone="Sheared DNA-46J23"						
/clone.lib="Sheared DNA"						
/note=Vector: pUC19; Site_1: SmaI; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (approx 2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barell, Oxford University						

JOURNAL COMMENT		Unpublished (2001) Contact: L. David Sibley WashU Plasmodium EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (sibley@borcim.wustl.edu), Washington University Seq primer: -400P from Gibco High quality sequence stop: 431.	
FEATURES		Location/Qualifiers	
source		1..562	
		/organism="Plasmodium falciparum"	
		/db_xref="taxon:5833"	
		/clone_lib="Plasmodium falciparum 3D7 asexual cDNA"	
		/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"	
		/note="Vector: pBluescript SK plus; Site.1: EcoRI; Site.2: XhoI; Library was constructed by Debopam Chakrabarti. Total RNA samples were isolated from mixed stage saponin(0.1%) lysed P. falciparum 3D7 infected erythrocytes by the acidic guanidinium-phenol chloroform method. The poly A+ RNA was isolated by the polyAT-tract mRNA isolation system (Promega, WI) using streptavidin Magnosphere particles. Directional cDNA libraries were constructed by oligo d(T) priming of poly(A)+ RNA (5mg) into EcoRI and XhoI sites of 1 ZapII vector using the Zap cDNA synthesis kit (Stratagene, CA). The average size of the cDNA inserts in the library was between 1.0 and 1.5kb. Clones were mass excised using the ExAssist helper phage (Stratagene), the phagemids were precipitated with PEG 8000 and extracted with phenol/chloroform. Phagemid DNA was electroporated into DH10B cells."	
BASE COUNT		290 a 35 c 47 g 190 t	
ORIGIN		Query Match 2.7%; Score 80; DB 13; Length 562; Best Local Similarity 47.4%; Pred. No. 1.9e-05; Matches 239; Conservative 0; Mismatches 265; Indels 0; Gaps 0;	
Qy	2023	TCATTAATAGACAGGCATCAATTATTAATATATGATGATTTAGATGAA	2082
Db	41	TTATTAACAATAAACAATTATAGATTTATTTCCACCGGATTTGATTCATATTAAC	100
Qy	2083	AGTAGAAGTATATCTTCAATAAAACCGCTGAATATCAAGAAGAAATGGATATGTTAAAA	2142
Db	101	TTTTTAATTAATCTTTAGATAGAAAAGATAAGATAATGATAAAGATAATGATAACAAT	160
Qy	2143	GATCTTAAACATCAATCAAAATAGATGCCAAATCCGAGTCAGTATTTCTATCCCAAT	2202
Db	161	AATGATAACCATATTTGATAACAATAATGATTAATTAATGAGATATTTGGACGGTGAAT	220
Qy	2203	ACATTAACGGCTTAGTGAAAAAACTAAACAATAATTAATGATTAATAATAAATAATTC	2262
Db	221	AAATATAAGAATTATAAAATTTATATAGAAATTTTAAAGATGATAAAAAATAATAATA	280
Qy	2263	GTTATTTCTACATGTTGATAAGATCATCTTACACAGACATCAAAAAGAAATATGT	2322
Db	281	AAAAATAATAATCAAAATGAAGAAAAAATTCATTTATACATAAATAATATCATATGT	340
Qy	2323	GCTTCTATCATAGCACAAGTGAATATTTTACTAAATAATGACATCTCATATTACACG	2382
Db	341	GATGATATATATATCAACATATTTATATTTTAAATAATGAAAACAACAATCATACA	400
Qy	2383	AGTATAGACTAAATAAAACCTAGGCACATTTAAGTAAATATTAATAATAGTCAGTTA	2442
Db	401	TTTGATAACCTTCATATACATATTGAGATGTTTAAAGATATTTTAAATTTTATTCATATA	460
Qy	2443	AATCTAAATTTGGAATACATCATTTTTCATATCATGACAGCCTATTCGTTAAAAATGAC	2502
BASE COUNT		213 a 35 c 10 g 383 t	
ORIGIN		Press, 1999)."	
Query Match		2.7%; Score 81.8; DB 17; Length 641;	
Best Local Similarity		47.1%; Pred. No. 8.4e-06;	
Matches 283; Conservative 0; Mismatches 317; Indels 1; Gaps 1;			
Qy	2029	AATAGACAGGCATCAATATATATATGACAAATTTGATGATTTAGATGAAAGTAGA	2088
Db	608	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	549
Qy	2089	AAATATATCTTCAATAAAACCGCTGAATATCAAGAAGAAATCGATATGTTAAAGATCTT	2148
Db	548	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	489
Qy	2149	AAATCATTCACAAATAAAGATGCCAAATTCGCAGTCAGTATTTCTATCCCAATACATTA	2208
Db	488	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	429
Qy	2209	AAAGCGTTAGTGAAGAACTAAACAATATATTGAATATAATAAAATATATTCGTTAT	2268
Db	428	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	369
Qy	2269	ATTCTACATGTTGATAGATCATCTTACACCAGACATCAAAAAGAAATATTCGGCTTC	2328
Db	368	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	309
Qy	2329	TATCATAAGCACCAAGTGAATATTTTACTAAATAATGACATCTCATATTACAGGATAT	2388
Db	308	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	249
Qy	2389	AGACTATAAAACTGAGGCACATTTAAGTAATATAATAAAATTAAGTCAGTTAAATCTA	2448
Db	248	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	189
Qy	2449	AATGTGAATACATCATTTTGTGATATCATGACAGCCTATTCGTTAAAAATGACAGCAT	2508
Db	188	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	129
Qy	2509	GCTTATATGAAAAATATGATGCGGCATGAATTTCTCAGCATTAACACATGATGGATC	2568
Db	128	ATAATAATAATAATAATAATGATGATGATGATGATGATGATGATGATGATGATGATG	70
Qy	2569	GAGAAATCAATCGCATCCACCATTTTAAAGAGCTGATTAATAAACTATTTTAATGACAAT	2628
Db	69	GTGTGATATATACGAATATATATAGTAAAGCTGCTGGTTAATGCTGATTCATCCAGAAA	10
Qy	2629	G 2629	
Db	9	G 9	
RESULT 5		BI815821	
LOCUS		PFE00a33a12.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium	
DEFINITION		falciparum cDNA 5' similar to TR:096216 096216 PREDICTED MEMBRANE ASSOCIATED PROTEIN. ; mRNA sequence.	
ACCESSION		BI815821	
VERSION		BI815821.1	
KEYWORDS		EST.	
SOURCE		malaria parasite P. falciparum.	
ORGANISM		Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.	
REFERENCE		1 (bases 1 to 562)	
AUTHORS		Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D.,	
		Matta, M., Hillier, L., Martin, J., Wyllie, R., Dente, M., Theising, B.,	
		Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,	
		Tsagareishvili, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,	
		Maquire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D.,	
		Waterson, R., Willson, R. and Sibley, D.	
TITLE		WashU Plasmodium EST Project	

Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE AUTHORS

1 (bases 1 to 1169)
Souciet J.L., Aigle M., Artiguenave F., Blandin G.,
Bolotin-Fukuhara M., Bon E., Brottier P., Casaregola S.,
de-Montigny J., Dujon B., Durrens P., Lepingle A., Llorente B.,
Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Potier S.,
Saurin W., Tekala F., Toffano-Nioche C., Wesolowski-Louvel M.,
Wincker P. and Weissenbach J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711
11152876
2 (bases 1 to 1169)
Casaregola S., Lepingle A., Bon E., Neuveglise C., Nguyen H.,
Artiguenave F., Wincker P. and Gaillardin C.
Genomic exploration of the hemiascomycetous yeasts: 7.
Saccharomyces servazzii
FEBS Lett. 487 (1), 47-51 (2000)
20584717
11152882
3 (bases 1 to 1169)
Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

COMMENT

1. .1169
/organism="Saccharomyces servazzii"
/strain="CBS 4311"
/db_xref="taxon:27293"
/clone="AT0AA005D05"
/clone_lib="AT0AA"
/note="end : T3"
BASE COUNT 375 a 101 c 57 g 545 t 91 others
ORIGIN

FEATURES

source

Query Match 2.6%; Score 77.4; DB 17; Length 1169;
Best Local Similarity 43.5%; Pred. No. 5.5e-05;
Matches 249; Conservative 26; Mismatches 294; Indels 4; Gaps 2;
Qy 1926 TAATAAATCTGCTATACCGCGTATTACATGGTGATACACATCCATTAGAACTCGG 1985
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1140 TAAATTAATATATATAAATAATAATAATATAATAAATAAATAAATAAATAA 1081
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 1986 CATCAAGAAGAACCACTTTGTTGTAGTCAATCAGTCATTAATACAGACGAGGCA 2045
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1080 TAAATAAATATAAATAAATAATAAATAAATAAATAAATAAATAAATAA 1021
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2046 TTATTATATTTATGACAAATTTGATTTAGATGAAAGTAGAAGTAGTATCTTCAATA 2105
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1020 AAATATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 961
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2106 AACCGCTGAATATCAAGAAGAAATGGATATGTTAAAGATCTTAAACATTCATCA 2165
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 960 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 901
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2166 AGATGCCAAATCCGAGTCAGTATTTCTATCCCAATACATTAACGCTTAGTCAAAA 2225
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 900 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 844
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2226 ACTAACCAATATTATTGAATATAAATAAATAAATAAATAAATAAATAAATAA 2285
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

Db 843 AAANAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 784
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2286 GAATCATCTTTACACAGACATCAAAAAGAAATATTGGCTTCTATCATACCAAGT 2345
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 783 TATTAT-ATAAANAATAAATAAATAAATAAATAAATAAATAAATAAATA 725
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2346 GAATATTTTACTAATAATACATCTCATATTACACAGAGTAAATAGACATTA 2405
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 724 AAAAAATATATTTATATTTTATATATATATATATATATATATATATATA 565
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2406 GGCACATTTAAGTAAATTAATAAATTAAGTCAGTAAATCTAAATTCGAATACAT 2465
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 664 TATATATATNTTAAATTAATAAATAAATAAATAAATAAATAAATAAATA 605
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Qy 2466 TTTTCATATATCATGACACGCTTATTCGTTAAAAA 2498
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 604 AAATANATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 572
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 10

CNS04NOJ CNS04NOJ 994 bp DNA linear GSS 21-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 122P04 of library G from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION.

VERSION AL298972.1 GI:8037822

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis.

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE

AUTHORS

1 (bases 1 to 994)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.

TITLE

Tetraodon nigroviridis DNA sequence
Unpublished

JOURNAL

REFERENCE

2 (bases 1 to 994)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis

JOURNAL

REFERENCE

3 (bases 1 to 994)
Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-and sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES

source

1. .994
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="122P04"
/clone_lib="G"
/note="Genoscope sequence ID : C08G122DH02LPI-end : T7"
BASE COUNT 543 a 49 c 36 g 194 t 172 others
ORIGIN

Query Match

2.6%; Score 76.8; DB 17; Length 994;

Best Local Similarity

37.8%; Pred. No. 7.3e-05;

Matches

318; Conservative 64; Mismatches 448; Indels 11; Gaps 2;

Qy

1664 AAGAATTTTAAAGATAAAGCGTAGCTGTGTGTATACCACTAATAAGAAACGTCATC 1723
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

[illegible]

```

McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model.
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants: project phase 2
Published (2000)
Contact: Charles Hauser
DCMB Box 91000
Durham, NC 27708-1000
Tel.: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
source      .973
            /organism="Chlamydomonas reinhardtii"
            /strain="CC-1690 wild type mt+ 2lgr"
            /db_xref="taxon:3055"
            /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
            II"
            /note="vector: pbluescript II SK-; Site_1: EcoRI; Site_2:
            XhoI; This library, constructed by John Davies and Jeffrey
            McDermott, combines cDNAs from CC-1690 cells grown to
            mid-log phase in TAP (acetate-containing) medium in the
            light, TAP medium in the dark, HS (minimal) medium in
            ambient levels of CO2 and HS medium bubbled with 5% CO2.
            PolyA mRNA was purified from each sample, pooled and cDNA
            synthesized. The cDNA was directionally cloned into lambda
            ZAP II (Stratagene). In the EcoRI (5') and XhoI (3') sites.
            pBluescript II SK- plasmids were excised from the lambda
            ZAP clones by superinfection with ExAssist (Stratagene)
            phage. The library was normalized using method 4 described
            in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT   189 a    51 c     89 g    642 t       2 others
ORIGIN
Query Match          2.6%; Score 76.4; DB 12; Length 973;
Best Local Similarity 47.3%; Pred. No. 8.7e-05;
Matches 262; Conservative 0; Mismatches 291; Indels 1; Gaps 1;
QY 1972 ATTAGAARACTCGGATTCAAAGAAACCAATTTTGTGTAGTCGAATTAATAAT 2031
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 838 ATTAACAAAATATATAAAAAATAGAAATAATAATTAATTAACAAAATTAATAAA 779
QY 2032 AGACAAGGCATCATATTTATTAATTAGACAAATTTTGATGGATGAAGTGAAGA 2091
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 778 ATAATAATATTTANTAAATTAATAAATAAACAAAGAAAGATATAAAAAATAAA 719
QY 2092 TATATCTTCAATTAACACCCTGAATATCAAGAGAANAATGGATATGTATTAAGCATCTTAAA 2151
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 718 AAATAAAAAAAAAAGTAATCCAACAAACATTAATAATAATAAAAAATAAAAAA 659
QY 2152 CTCATTTCAAAATAAAGATGCCAAATCGCAGTCAGTATTTTTCTATCCCATAACATTAAC 2211
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 658 ATAAATTAACAAAAAAGCAACAAAAAATAAATAAACATAATAAAAAATTAACAAA 599
QY 2212 GCCTTAGTGAACCAACTAAACAATATATTGGAATATAATAAATAATNTCGTTATTATT 2271
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 598 GAATTAAATATAATATAATAAATAAAAAAATAAAAAAATAAAAAATAAAAAAT 539
QY 2272 CTACATGTTGAATGAATCATCTTACACGAGCATCAAAACAGAAATATTTGGCTTTCTAT 2331
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 538 AACAAAAAATAAAAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAT 479
QY 2332 CATTAAGCACCAAGTGAATATTTTACTAAATAATGACATCTCATATPACAGAGTAATAGA 2391
Ta -AATTTACAAAAATAAATAAATAAACAATAAAAAAATAATACATAAAAAAACAACATAAAT 420
QY 2392 CTAATAAAAACTGAGGCACATTTTAAGTAATATTAAATAAAATTAAGTCAGTTAAATCTAAAT 2451
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 419 AAAAATAAAAAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATTCAC 360
QY 2452 TGTCATATCATCATTTTTTGTGAATCATGACAGCCCTATTCGTTAAAAATGACAGCTATGCT 2511
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY	2215	TTAGTGA	AAAAA	CTAAACCAATATTATTGAATATAATAAAAAATATATTGCTTA	TATTTCTTA	TTATTTCTTA	2274
Dd	329	AAAATTA	AAAAA	TAATAAAAAA	TAATAAAAAA	TGTTAAAGA	388
QY	2275	CATGTTG	AAGAATCATCTTACAC	CAGACATCAAA	AAAAAGAAATATTGGC	TTTTCTTCATCAT	2334
Dd	389	ATAAATA	AAAAA	AAAAA	AAAAA	AAAAA	448
QY	2335	AAGCACCA	ACTGAATATTTTACT	TAAATAATGACATCTCATAT	TACACGAGTA	ATAAGACTA	2394
Dd	449	AAATAAA	AGAAAATAAA	AAAAATAAATAATATA	TATAAATAA	GAGATAA	508
QY	2395	ATAAAACT	GAGCGACATTTTAAGTA	TATATATAAATTAAG---	TCAGTTAAATCT	TAAT	2455
Dd	509	AAATATATA	AAAAAATGAAATTA	AAAAAATTAATAAAT	AAAAAGATAA	TAAGATATAGAT	568
QY	2452	TGTGAAT	FACATCATTTTTTGAT	ATATCATGACGCCTATTCGTT	AAAAAA	2498	
Dd	569	TGTAAAA	AAATACACATATAATA	TATAATAAGAA	AAAAA	TAATAAANA	615
RESULT	13						
LOCUS	BH391984						
DEFINITION	Ag-ND-138G17.TR ND-TAM Anopheles gambiae genomic clone AG-ND-11-DEC-						
ACCESSION	BH391984						
VERSION	BH391984.1						
KEYWORDS	GSS.						
SOURCE	African malaria mosquito.						
ORGANISM	Anopheles gambiae						
REFERENCE	Shetty,J., Matek,J., Koo,H., Collins,F., Gardner,M. and Loftus						
AUTHORS	Direct Submission of BAC-end sequences from Anopheles gambiae						
TITLE	Unpublished (2001)						
JOURNAL	Other GSSs: AG-ND-138G17.TF.1						
COMMENT	Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: bjloftus@tigr.org This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest. Seq primer: M13 Rev Class: BAC ends.						
FEATURES	source						
	1..832	Location/Qualifiers					
	/organism="Anopheles gambiae"						
	/strain="PEST"						
	/db_xref="taxon:7165"						
	/clone="AG-ND-138G17"						
	/clone_lib="ND-TAM"						
	/note="Vector: pECBAC1; Site_1: HindIII"						
BASE COUNT	425 a	71 c	59 g	277 t			
ORIGIN							
Query Match	2.6%;	Score 76;	DB 17;	Length 832;			
Best Local Similarity	46.1%;	Pred. No. 0.00011;					
Matches	294;	Conservative	0;	Mismatches	340;	Indels	4;
OV	1915	TTTAAACATCTT	ATAA	AAATCTGCTAAT	CCCGGTATTACAT	GGTGATAACACATCCATT	1970

ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source
1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR05N11"
/clone_lib="RPCI-98"
/note="end: TET3"
/note="7 c 28 g 289 t 146 others

BASE COUNT 631 a 7 c 28 g 289 t 146 others
ORIGIN

Query Match 2.6%; Score 76; DB 17; Length 1101;
Best Local Similarity 40.1%; Pred. No. 0.0001;
Matches 300; Conservative 40; Mismatches 405; Indels 4; Gaps 1;

QY 1898 TCAGTGAAGCTTGGAAATTTAAACATCTTTAAATAATCTGTATATACCGCGTATTACATG 1957
DB 145 TCATGAAAGCTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 204
QY 1958 GTGATAACACATCCATTAAGAACTCGGCATTCAGAAAGAAACCAATTTTGTGTAGTCA 2017
DB 205 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 264
QY 2018 ATCAGTCATTAAATAGACAGGCATCAATTTATATATATGACAAATTTGATGATTTAG 2077
DB 265 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 324
QY 2078 ATGAAAGTGAAGTATATCTTCAATAAACCCTGAATATCAAGAAGAAATGGATATGT 2137
DB 325 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 384
QY 2138 TAAAGATCTTAAATCAATTCATAAATAAGATGCCAAATCGCAGTCAGTATTTCTATC 2197
DB 385 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 444
QY 2198 CCAATACATTTAAACGGCTTAGTGAAACAACTAAACATATATTTGATATATAATAATA 2257
DB 445 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 504
QY 2258 TATTCGTTATATCTACATGTTGATAGAAATCATCTTACACAGACATCAAAAAAGAAA 2317
DB 505 ANAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 564
QY 2318 TATTCGCTTCTATCATAAGCAGCAGTGAATATTTTACTAAAT----AATGACATCTCA 2373
DB 565 TTTAAATTTTAAAWAAATTTTAAATAAAWATTTTAAATTTTAAATTTTAAATTTTAAAT 624
QY 2374 TATTACAGGAGTATAGACATTAATAAATCGAGGACATTTTAAAGTAAATATAATAATA 2433
DB 625 TTTTAAAWAATTTTATTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATA 684
QY 2434 AGTCAGTTTAAATCTAAATTTGTAATACATCATTTTGTGATAATCATGACAGCCTATCT 2493
DB 685 AATWAAATTTTAAATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 744
QY 2494 AAAAAATGACAGCTATGCTTATATGAAAAAATATGATCGCGCATGAATTTCTCAGCATTA 2553
DB 745 WATTTWTAATTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 804
QY 2554 ACACATGATTTGGATCGAGAAATCAATCGCATCCACCATTTTAAATAAGCTGATTAAC 2613
DB 805 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 864
QY 2614 TATTTTAAATGACATGCTTTAGAAAGTAT 2642
DB 865 ATTTTATTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 893

DB 150 TTTATCATTCGGTTAAATATACTATAATAATAATAATAATAATAATAATAATAATAATA 209
QY 1975 AAGAACTCGCATTCAGAAAGAAACCACTTTTGTGTAGTCAATCAGTCATTTAAATAGA 2034
DB 210 ATAAATACATCATCATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 269
QY 2035 CAAGGCATCAATTTATATAATTTATGACAAATTTGATGATTTAGATGAAGATAGAAAGTAT 2094
DB 270 AATAATAACATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 329
QY 2095 ATCTTCAATAAACCCTGATATATCAAGAGAAATGGATATCTTTAAAGATCTTAAACCTC 2154
DB 330 ATTATAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 389
QY 2155 ATTCAAAATAAAGATGCCAAAATCGCAGTCAGTATTTTCTATCCCAATACATTTAAACGGC 2214
DB 390 AATA 449
QY 2215 TTAGTGAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2274
DB 450 AATA 509
QY 2275 CATGTTGATAAGATCAT---CTTACACAGACATCAAAAAAGAAATATTTGGCTTTCTA 2330
DB 510 AATA 569
QY 2331 TCATAAGCCAAAGTGAATATTTTACTTAATAATGACATCTCATATTACACGAGTAAATAG 2390
DB 570 AAAAAAGATAATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 629
QY 2391 ACTAATAAATACTGAGGCACATTTAAGTAAATATAATAATAATAATAATAATAATAATA 2450
DB 630 TAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 689
QY 2451 TTGTAATACATCATTTTGTGATAATCATGACAGCCATTTGCTTTAAATAAGACAGCTATGC 2510
DB 690 TAATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 749
QY 2511 TTATATGAAAAATATGATCGCGCATGAATTTCTCAG 2548
DB 750 ATTATTTGGATTATGATATATAATAACATATCTG 787

RESULT 14
CNS0021J 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence IET3 end of BAC #
DEFINITION BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL061936
VERSION AL061936.1 GI:4940214
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 91 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>
The BDGP BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammox in Pletier de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 00:35:34 ; Search time 118 Seconds
(without alignments)
1089.719 Million cell updates/sec

Title: US-09-842-484a-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGNIPIVNRKFIINSITL 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_101002:*			
1:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*		
2:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*		
3:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*		
4:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*		
5:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*		
6:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*		
7:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*		
8:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*		
9:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*		
10:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*		
11:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*		
12:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*		
13:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*		
14:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*		
15:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*		
16:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*		
17:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*		
18:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*		
19:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*		
20:	/SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*		
21:	/SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*		
22:	/SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*		
23:	/SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5089	100.0	965	21	AA196213
2	5089	100.0	965	23	AA196213
3	5067	99.6	965	23	AA196213
4	4486.5	88.2	972	20	AA196212
5	4486.5	88.2	972	21	AA196212
6	3079	60.5	702	21	AA196212
7	292.5	5.7	327	23	AA196212
8	276.5	5.4	706	22	AA196212
9	276.5	5.4	715	22	AA196212
10	253	5.0	321	23	AA196212

11	251	4.9	298	22	AA196213	Putative glycosylt
12	246	4.8	322	21	AA196213	CpsII protein whic
13	238.5	4.7	324	18	AA196213	S.thermophilus exo
14	238.5	4.5	339	21	AA196213	S.thermophilus exo
15	230.5	4.5	339	21	AA196213	Enzyme EPS5 involv
16	230.5	4.5	339	21	AA196213	Enzyme EPS5 involv
17	228.5	4.5	322	21	AA196213	Amino acid sequenc
18	228.5	4.5	322	21	AA196213	Enzyme EPS7 which
19	227	4.5	322	21	AA196213	Amino acid sequenc
20	225.5	4.4	316	21	AA196213	Staphylococcus epi
21	225.5	4.4	316	21	AA196213	Enzyme EPS8 which
22	223.5	4.4	278	21	AA196213	Amino acid sequenc
23	223.5	4.4	674	19	AA196213	CpsII protein whic
24	223.5	4.4	674	23	AA196213	Streptococcus pneu
25	217.5	4.3	332	21	AA196213	S. pneumoniae Sp11
26	217.5	4.3	332	21	AA196213	Cps2J protein whic
27	217	4.3	330	21	AA196213	Cps2K protein whic
28	217	4.3	330	21	AA196213	Enzyme EPS6 involv
29	216	4.2	336	21	AA196213	Amino acid sequenc
30	216	4.2	336	21	AA196213	Enzyme EPSM involv
31	215.5	4.2	346	23	AA196213	Amino acid sequenc
32	213.5	4.2	327	21	AA196213	Neisseria meningit
33	213.5	4.2	327	21	AA196213	Enzyme EPSJ involv
34	210	4.1	348	17	AA196213	Amino acid sequenc
35	210	4.1	348	18	AA196213	N. gonorrhoeae gly
36	209	4.1	337	17	AA196213	Neisseria polysylc
37	209	4.1	337	18	AA196213	N. gonorrhoeae gly
38	208.5	4.1	270	21	AA196213	Lipo-oligosacchari
39	207	4.1	301	21	AA196213	Campylobacter jeju
40	206	4.0	328	21	AA196213	Campylobacter jeju
41	206	4.0	329	22	AA196213	Streptococcus pneu
42	205	4.0	389	21	AA196213	EpsM. Lactococcus
43	203	4.0	332	22	AA196213	Campylobacter jeju
44	202	4.0	302	21	AA196213	EpsN. Lactococcus
45	201.5	4.0	324	23	AA196213	Campylobacter jeju
						Lactococcus lactis

ALIGNMENTS

RESULT 1	
AA196213	
ID	AA196213 standard; Protein; 965 AA.
XX	
AC	AA196213;
XX	
DF	17-AUG-2000 (first entry)
XX	
DE	P. multocida chondroitin synthase.
XX	
KW	Chondroitin synthase; CS; enzyme; hyaluronic acid; ulcer;
KW	tissue abrasion; viscoelastic replacement; bioadhesive.
XX	
OS	Pasteurella multocida.
XX	
PN	WO200027437-A2.
XX	
PD	18-MAY-2000.
XX	
PF	10-NOV-1999; 99WO-US26501.
XX	
PR	11-NOV-1998; 98US-0107929.
XX	
PR	01-APR-1999; 99US-0283402.
XX	
PA	(OKLA) UNIV OKLAHOMA STATE.
XX	
PI	DeAngelis PL;
XX	
DR	WPI; 2000-376319/32.
XX	
PT	N-PSDB; AAA27449.
XX	
PT	Novel method for the enzymatic transfer of sugar molecules to an
XX	acceptor, useful for synthesis of e.g. polysaccharide bioadhesives or

drug delivery systems, including hybrid molecules

Claim 15; Page 85; 86pp; English.

The present sequence is the Pasteurella multocida chondroitin synthase pMCS. pMCS catalyses glycosaminoglycan polymerisation to produce chondroitin: a linear polysaccharide which has viscoelastic properties which makes it useful for a number of applications. Chondroitin can be used with hyaluronic acid (HA) to coat medical devices e.g. catheters and sensors to reduce tissue abrasion. In addition, they can be used as bioadhesives for haemostatic sealing and healing of wounds and surgical incisions; and as biomaterials that provide sustained delivery of encapsulated drugs, to wounds, ulcers, injuries or surgical sites.

Sequence 965 AA;

Query Match	100.0%;	Score 5089;	DB 21;	Length 965;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 965;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNTLSQAIKAYNSNDYELALKEFSAETVGRKIVFEQIICKCKELSTNSVSEDKNSV	60	
Db	1	MNTLSQAIKAYNSNDYELALKEFSAETVGRKIVFEQIICKCKELSTNSVSEDKNSV	60	
QY	61	CDSSLDIATOLLNSVKKLTLSSEKNSLKNWKSITGKSENAETRKVELVPKDPKDL	120	
Db	61	CDSSLDIATOLLNSVKKLTLSSEKNSLKNWKSITGKSENAETRKVELVPKDPKDL	120	
QY	121	VLAPLPHVNDFTWYKRNKSLGKIPVKNIGLSIIPTENRSRIILDTIACLVNOKTNY	180	
Db	121	VLAPLPHVNDFTWYKRNKSLGKIPVKNIGLSIIPTENRSRIILDTIACLVNOKTNY	180	
QY	181	PFEVVADGSKENLTIIVQYKQKDIKIVYRQKDYQICAVRNGLRTAKYDFVSILD	240	
Db	181	PFEVVADGSKENLTIIVQYKQKDIKIVYRQKDYQICAVRNGLRTAKYDFVSILD	240	
QY	241	CDMAPOQLVHSHYLTLELNDIVLIGPKYVDTHNTAEQFLNDPYLESLPETATNN	300	
Db	241	CDMAPOQLVHSHYLTLELNDIVLIGPKYVDTHNTAEQFLNDPYLESLPETATNN	300	
QY	301	PSITSKGNISLDWRLEHFKFTDNLRLCDSPPRYVAGNVAFSEKWLKNVGHVDFEENHWG	360	
Db	301	PSITSKGNISLDWRLEHFKFTDNLRLCDSPPRYVAGNVAFSEKWLKNVGHVDFEENHWG	360	
QY	361	GEDVEFSYRLFAKGFRRVIDGGMAIHQEPGKENETEREAGKSTILKIVKEKVPYIYRK	420	
Db	361	GEDVEFSYRLFAKGFRRVIDGGMAIHQEPGKENETEREAGKSTILKIVKEKVPYIYRK	420	
QY	421	LLPIEDSHIHRPLVSIYIPAYNCANYIQRVCDSALNQTVDLEVCICNDGSTNTLEVI	480	
Db	421	LLPIEDSHIHRPLVSIYIPAYNCANYIQRVCDSALNQTVDLEVCICNDGSTNTLEVI	480	
QY	481	NKLYGNPRVRIMSPNGGSIASASNAVSFAKGYVIGOLDDSDYLEPDAVELCLKEFLKD	540	
Db	481	NKLYGNPRVRIMSPNGGSIASASNAVSFAKGYVIGOLDDSDYLEPDAVELCLKEFLKD	540	
QY	541	KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTATIAHFRMFTIRAWHLTDGFNEN	600	
Db	541	KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTATIAHFRMFTIRAWHLTDGFNEN	600	
QY	601	IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHFVVVQNSLNROG	660	
Db	601	IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHFVVVQNSLNROG	660	
QY	661	INYNYDKFDDDESRYIFNKTAEYQOERMDMLKDLKLIQNKDAKIAVSIYPNTLNGLV	720	
Db	661	INYNYDKFDDDESRYIFNKTAEYQOERMDMLKDLKLIQNKDAKIAVSIYPNTLNGLV	720	
QY	721	KKLNNIIEYNKNFVILHVDKNHLPDPIKKEILAPYHKHQVNILLNNDISYITSNRLIK	780	
Db	721	KKLNNIIEYNKNFVILHVDKNHLPDPIKKEILAPYHKHQVNILLNNDISYITSNRLIK	780	
QY	781	TEAHLNSINKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWEIKIN	840	

Db	781	TEAHLNSINKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWEIKIN	840	
QY	841	AHPPEKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQSIDSVPEY	900	
Db	841	AHPPEKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQSIDSVPEY	900	
QY	901	NTEDWFOFALLILEKKTGHVFNKTSTITYMPWERKLOWNTNQIOSAKKGENIPVKNKFI	960	
Db	901	NTEDWFOFALLILEKKTGHVFNKTSTITYMPWERKLOWNTNQIOSAKKGENIPVKNKFI	960	
QY	961	NSITL 965		
Db	961	NSITL 965		

RESULT 2

AAM47335
ID AAM47335 standard; Protein; 965 AA.

XX

AC AAM47335;

DT 22-FEB-2002 (first entry)

XX Pasteurella multocida chondroitin synthase #1.

XX Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;

KW eye application; joint application; moisturiser; drug delivery;

KW wound dressing; biocompatible film.

XX Pasteurella multocida.

OS WO200180810-A2.

PN 01-NOV-2001.

PD 25-APR-2001; 2001WO-US13395.

XX 25-APR-2000; 2000US-199538P.

PR (DANG/) DE ANGELIS P L.

XX De Angelis PL;

PI WPI; 2002-049237/06.

XX N-PSDB; ABAU5097.

DR New chondroitin synthase gene obtained from Pasteurella multocida,
XX useful as hyaluronan polysaccharide substitute in medical or cosmetic
XX applications, e.g. for eye or joint applications, for moisturizer or
XX wound dressings

PS Claim 3; Page 119; 125pp; English.

XX The present invention relates to the coding sequence of the Pasteurella
XX multocida chondroitin synthase. A chondroitin polysaccharide may be used
XX as a hyaluronan polysaccharide substitute in medical or cosmetic
XX applications, for example in eye or joint applications, for moisturiser
XX or wound dressings. The enzyme may be used in covalently coupling
XX specific drugs, proteins or toxins to the structurally modified
XX chondroitin for general or targeted drug delivery or radiological
XX procedures, covalently cross linking the hyaluronan acid itself or to
XX other supports to achieve a gel or other three dimensional biomaterial
XX with stronger physical properties, and covalently linking hyaluronan acid
XX to a surface to create a biocompatible film or monolayer. The present
XX sequence is one version of the protein of the invention.

XX Sequence 965 AA;

Query Match

Best Local Similarity 100.0%;

Matches 965; Conservative 0;

Mismatches 0;

Indels 0;

Gaps 0;

Score 5089; DB 23; Length 965;

Pred. No. 0;

```
QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
DB 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
QY 61 CDSLSLDIATQLLSNVKLLTSLSEKNSLNKNKWSITGKKSENAEIRKVELVPKDPKDL 120
DB 61 CDSLSLDIATQLLSNVKLLTSLSEKNSLNKNKWSITGKKSENAEIRKVELVPKDPKDL 120
QY 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKIGLSIIPTFNRSRLDITLACLVNOKTNY 180
DB 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKIGLSIIPTFNRSRLDITLACLVNOKTNY 180
QY 181 PFEVVVADGSGKENLLTIVQKYEQKLDIKYVRQDKYQYLCVARNLGLRTAKYDFVSILD 240
DB 181 PFEVVVADGSGKENLLTIVQKYEQKLDIKYVRQDKYQYLCVARNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPVIIESLPETATNN 300
DB 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPVIIESLPETATNN 300
QY 301 PSITSKGNISLDRLEHFKKTDNLRLCDSPFYFVAGNVAFSKEWLNKVGWFEDEFNHWG 360
DB 301 PSITSKGNISLDRLEHFKKTDNLRLCDSPFYFVAGNVAFSKEWLNKVGWFEDEFNHWG 360
QY 361 GEDVEGYRLFAKGCFFRVIDGMAIHQEPGKRENETEREAGKSITLKIYKVKVPYIYRK 420
DB 361 GEDVEGYRLFAKGCFFRVIDGMAIHQEPGKRENETEREAGKSITLKIYKVKVPYIYRK 420
QY 421 LLPIEDSHIIRPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICDNGSDTNTLEVI 480
DB 421 LLPIEDSHIIRPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICDNGSDTNTLEVI 480
QY 481 NKLYGNPNRVRIMSKPNGGGLASNAVSAFAGYVIGQLSDDDVLEPDVAVELCLKEFLKD 540
DB 481 NKLYGNPNRVRIMSKPNGGGLASNAVSAFAGYVIGQLSDDDVLEPDVAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
DB 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKLVSEVGFHKLKICYNRVLHGDNTSIKKLGIOKKNHFVVVQSLNRQ 660
DB 601 IENAVDYDMFLKLVSEVGFHKLKICYNRVLHGDNTSIKKLGIOKKNHFVVVQSLNRQ 660
QY 661 INYNYDKFDLDESRRYIENKTAEYQEDMLKDLKIQNKDAKIAVSIFYPNTLGLV 720
DB 661 INYNYDKFDLDESRRYIENKTAEYQEDMLKDLKIQNKDAKIAVSIFYPNTLGLV 720
QY 721 KLLNNIIEYNKNIFFVIIHLVDKNHLPDIIKKEILAFYHKHQVNILLNNDISYYSNRLIK 780
DB 721 KLLNNIIEYNKNIFFVIIHLVDKNHLPDIIKKEILAFYHKHQVNILLNNDISYYSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWIEKIN 840
DB 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWIEKIN 840
QY 841 AHPPEFKLIKTIFYNDNDRSMNVKGSQGMFKYALPHELLTIKEIVITSCQSDTSYPEY 900
DB 841 AHPPEFKLIKTIFYNDNDRSMNVKGSQGMFKYALPHELLTIKEIVITSCQSDTSYPEY 900
QY 901 NTEIDWIFQFALLILEKKTGHVFNKTSITLYMPWERKLTQWNEQIQSAKKGENTPVNKFII 960
DB 901 NTEIDWIFQFALLILEKKTGHVFNKTSITLYMPWERKLTQWNEQIQSAKKGENTPVNKFII 960
QY 961 NSITL 965
DB 961 NSITL 965
```

RESULT 3

AA047336

ID AA047336 standard; Protein: 965 AA.

XX

```
AC AA047336;
XX 22-FEB-2002 (first entry)
XX Pasteurella multocida chondroitin synthase #2.
DE Chondroitin synthase; dermatan sulphate; chondroitin sulphate polymer;
KW eye application; joint application; moisturiser; drug delivery;
KW wound dressing; biocompatible film.
XX Pasteurella multocida.
OS WO200180810-A2.
XX PN 01-NOV-2001.
XX PD 25-APR-2001; 2001WO-US13395.
XX PF 25-APR-2000; 2000US-199538P.
XX PR (DANG/) DE ANGELIS P L.
XX PA De Angelis PL;
PI WPI; 2002-049237/06.
XX DR N-PSDB; ABA05098.
XX PT New chondroitin synthase gene obtained from Pasteurella multocida,
PT useful as hyaluronan polysaccharide substitute in medical or cosmetic
PT applications, e.g. for eye or joint applications, for moisturizer or
PT wound dressings.
XX Claim 3; Page 122; 125pp; English.
PS The present invention relates to the coding sequence of the Pasteurella
CC multocida chondroitin synthase. A chondroitin polysaccharide may be used
CC as a hyaluronan polysaccharide substitute in medical or cosmetic
CC applications, for example in eye or joint applications, for moisturiser
CC or wound dressings. The enzyme may be used in covalently coupling
CC specific drugs, proteins or toxins to the structurally modified
CC chondroitin for general or targeted drug delivery or radiological
CC procedures, covalently cross linking the hyaluronic acid itself or to
CC other supports to achieve a gel or other three dimensional biomaterial
CC with stronger physical properties, and covalently linking hyaluronic acid
CC to a surface to create a biocompatible film or monolayer. The present
CC sequence is one version of the protein of the invention.
XX Sequence 965 AA;
QY Query Match 99.6%; Score 5067; DB 23; Length 965;
DB Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
DB 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
QY 61 CDSLSLDIATQLLSNVKLLTSLSEKNSLNKNKWSITGKKSENAEIRKVELVPKDPKDL 120
DB 61 CDSLSLDIATQLLSNVKLLTSLSEKNSLNKNKWSITGKKSENAEIRKVELVPKDPKDL 120
QY 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKIGLSIIPTFNRSRLDITLACLVNOKTNY 180
DB 121 VLAPLPHVNDFTWYKRNKSLGKIPVKNKIGLSIIPTFNRSRLDITLACLVNOKTNY 180
QY 181 PFEVVVADGSGKENLLTIVQKYEQKLDIKYVRQDKYQYLCVARNLGLRTAKYDFVSILD 240
DB 181 PFEVVVADGSGKENLLTIVQKYEQKLDIKYVRQDKYQYLCVARNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPVIIESLPETATNN 300
DB 241 CDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTAEQFLNDPVIIESLPETATNN 300
```


Db 601 TDGFNEKIENAVDYDMFLKSEVGFHKLKICYNRVLHGDNTSIKLGIOKKNHFVVVN 660
Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAEOEEMDKLKLIONKDAKIAVSIFYP 713
Db 661 QSLNRQGITTYNDEFDDDESRYIFNKTAEOEEDILDKLKLIONKDAKIAVSIFYP 720
Qy 714 NTLNGLVKKLNIIEYNKNIIFVILHVDKNHLPDIKKEILAFYHKKHQNILLNNDISYY 773
Db 721 NTLNGLVKKLNIIEYNKNIIFVILHVDKNHLPDIKKEILAFYHKKHQNILLNNDISYY 780
Qy 774 TSNRLIKTEAHLNSLNKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAUTH 833
Db 781 TSNRLIKTEAHLNSLNKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAUTH 840
Qy 834 DWIEKINAHPPFKLIKITYFNDNLDLSNMVKGASOGFMKVALPHELLTIIKEVITSQCS 893
Db 841 DWIEKINAHPPFKLIKITYFNDNLDLSNMVKGASOGFMVYALAHELLTIIKEVITSQCS 900
Qy 894 IDSYPEYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPWPERKQLQWNTNEQIOSAKKGENI 953
Db 901 IDSYPEYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPWPERKQLQWNTNEQIESAKKGENI 960
Qy 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972
RESULT 5
ID AAY43099 standard; Protein; 972 AA.
XX AAY43099;
XX AC AAY43099;
XX AD AAY43099;
DT 01-FEB-2000 (first entry)
XX P. multocida hyaluronate synthase (PmHAS) amino acid sequence.
DE Hyaluronate synthase; PmHAS; hyaluronan; hyaluronic acid; HA; cosmetic;
KW drug delivery; angiogenesis; wound healing; capsule synthesis;
KW fowl cholera; shipping fever.
XX Pasturella multocida.
OS
PN W09951265-A1.
PN 14-OCT-1999.
PF 01-APR-1999; 99WO-US07289.
PR 02-APR-1998; 98US-0080414.
PR 26-OCT-1998; 98US-0178851.
XX (OKLA) UNIV OKLAHOMA.
XX
XX Deangelis P;
XX
XX WPI: 2000-013032/01.
DR N-PSDB; AAZ35589.
XX
XX New isolated hyaluronate synthase nucleic acids, used for the
PT production of hyaluronic acid, for developing antibiotics and vaccines
PT and for diagnostic applications -
XX
XX Claim 70; Fig 20; 12lpp; English.
XX
XX This is the Pasturella multocida hyaluronate synthase (PmHAS) amino acid
CC sequence. Hyaluronic acid (HA) or hyaluronan, is a polysaccharide that
CC serves both structural and recognition roles in higher animals. Bacteria
CC produce extracellular capsules of HA which mimic their host HA and aid
CC escape from a host immune response. The invention includes a vector
CC containing the PmHAS nucleotide sequence which can be used to express
CC PmHAS in a foreign host. The HS nucleic acids can be used for the

CC production of HA. Also, specific changes to the HS coding sequence can
CC result in the production of HA having a modified size distribution or
CC structural configuration and functional properties. The HA products can
CC be used in e.g. drug delivery, angiogenesis and wound healing.
CC stabilisation of recombinant proteins and in cosmetics. The HS nucleic
CC acids can also be used to develop agents to block capsule synthesis by
CC pathogens and act as antibiotics. The avirulent P. multocida strains can
CC be used as vaccines for fowl cholera or shipping fever.
XX
Qy Sequence 972 AA;
Query Match 88.2%; Score 4486.5; DB 21; Length 972;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;
Qy 1 MNTLSQALKAYNSNDYELALKLFKSAETYGRIKIVFEFOIICKCKEL- - - STNS- - - YVS 53
Db 1 MNTLSQALKAYNSNDYELALKLFKSAETYGRIKIVFEFOIICKCKELSAHPSVNSAHSVN 60
Qy 54 BDKNSVCDSSLDIATQLLLSNVKKLTLSSEKNSLKNKWSITGKKSSENAIRKVELVP 113
Db 61 KEKVNVCDSPLDIATQLLLSNVKKLVLSDEKNTLKNKWLTTKKSSENAEVRVALVP 120
Qy 114 KDFPKDLVLAPLDHVNDFTWYKRNKKSGLGTPVKNKNTGLSIIITFTNRSRLDITLACL 173
Db 121 KDFPKDLVLAPLDHVNDFTWYKRNKKSGLGTPVKNKNTGLSIIITFTNRSRLDITLACL 180
Qy 174 VNQKTYPEVVVADGSGKENLLTIVQYEOKLDIKYVRQKDYGYQLCAVRNLGRTAKY 233
Db 181 VNQKTYPEVVIVTDGSGQEDLSPIIROYENKLDIRYVRQKDNQFOASARNMGLAKY 240
Qy 234 DFVSLDCDMPAQQWLWHSYLTLELNDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFGLLDCDMPANPLWHSYVAELLEDDDLTIIGPRKYIDTQHIDPKDFLNNASLESPL 300
Qy 294 ETATNNPSITSKGNISLDWRLHFKTKDNLRLCDSPPRYFVAGNVAFSKEWLNVKWPFD 353
Db 301 EVKTNNSVAAKGEVTSVLDWRLHFKTKDNLRLCDSPPRYFVAGNVAFSKWLNKSGFPD 360
Qy 354 EEFNHWGGEDVEFGYRLFAKCGFFRVIDGMAIHQEPCKENETREACKSITLAKIVKPK 413
Db 361 EEFNHWGGEDVEFGYRLFRYGSFFKTIIDGIMAYHQPCKENETDREAGKNTLDMIREK 420
Qy 414 VPYIYRKLPIEDSHIRIPLVSIYPAYNCANYIQRVCVDSALNQTVDVLEVCINDGST 473
Db 421 VPYIYRKLPIEDSHIRIPLVSIYPAYNCANYIQRVCVDSALNQTVDVLEVCINDGST 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGSIASASNAAYSAFAGYYIGQLDSDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGSIASASNAAYSAFAGYYIGQLDSDDYLEPDAVELC 540
Qy 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTTRAWHL 593
Db 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTTRAWHL 600
Qy 594 TDGFNEKIENAVDYDMFLKSEVGFHKLKICYNRVLHGDNTSIKLGIOKKNHFVVVN 653
Db 601 TDGFNEKIENAVDYDMFLKSEVGFHKLKICYNRVLHGDNTSIKLGIOKKNHFVVVN 660
Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAEOEEMDKLKLIONKDAKIAVSIFYP 713
Db 661 QSLNRQGITTYNDEFDDDESRYIFNKTAEOEEDILDKLKLIONKDAKIAVSIFYP 720
Qy 714 NTLNGLVKKLNIIEYNKNIIFVILHVDKNHLPDIKKEILAFYHKKHQNILLNNDISYY 773
Db 721 NTLNGLVKKLNIIEYNKNIIFVILHVDKNHLPDIKKEILAFYHKKHQNILLNNDISYY 780
Qy 774 TSNRLIKTEAHLNSLNKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAUTH 833
Db 781 TSNRLIKTEAHLNSLNKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAUTH 840
Qy 834 DWIEKINAHPPFKLIKITYFNDNLDLSNMVKGASOGFMKVALPHELLTIIKEVITSQCS 893
Db 841 DWIEKINAHPPFKLIKITYFNDNLDLSNMVKGASOGFMVYALAHELLTIIKEVITSQCS 900

841 DWTEKINAHPPFKLIKTYFNDNLKSMNVKGASQGMFTYALAHLLTIIKEVITSOS 900
894 IDSVPYNTEDWFOALLILEKTKGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENI 953
901 IDSVPYNTEDWFOALLILEKTKGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENI 960
954 PVNKFIINSITL 965
961 PVNKFIINSITL 972
RESULT 6
ID AAY96212 standard; Protein; 702 AA.
XX
AC AAY96212;
XX
DT 17-AUG-2000 (first entry)
XX
DE P. multocida hyaluronic acid synthase-D.
XX
KW Hyaluronic acid synthase-D; HAS-D; enzyme; ophthalmic surgery;
KW cataract; arthritis; ulcer; tissue abrasion; bioadhesive;
KW viscoelastic replacement; hyaluronic acid production.
XX
OS Pasteurella multocida.
XX
FH Key Location/Qualifiers
FT Misc-difference 450 /note= "encoded by AACtGT"
XX
XX WO200027437-A2.
XX
XX 18-MAY-2000.
XX
XX 10-NOV-1999; 99WO-US26501.
XX
XX 11-NOV-1998; 98US-0107929.
XX 01-APR-1999; 99US-0283402.
XX
XX (OKLA) UNIV OKLAHOMA STATE.
XX
XX DeAngelis PL;
XX
XX WPI: 2000-376319/32.
XX N-PSDB; AAA27448.
XX
XX Novel method for the enzymatic transfer of sugar molecules to an
XX acceptor, useful for synthesis of e.g. polysaccharide bioadhesives or
XX drug delivery systems, including hybrid molecules -
XX
XX Claim 11; Page 83; 86pp; English.
XX
XX The present sequence is the soluble recombinant form of Pasteurella
XX multocida hyaluronic acid synthase-D, PmHAS-D. This sequence encompasses
XX residues 1 to 703 of the 972 residues of the native PmHAS. PmHAS-D
XX catalyses glycosaminoglycan polymerisation to produce hyaluronic acid,
XX HA; a linear polysaccharide. HA has viscoelastic properties which makes
XX it useful for a number of applications. HA can be used during ophthalmic
XX surgery as a viscoelastic replacement for the vitreous humour e.g. during
XX implantation of intraocular lenses in cataract patients. HA injections
XX directly into joints is also used to alleviate pain associated with
XX arthritis. HA can also be used to coat medical devices e.g. catheters and
XX sensors to reduce tissue abrasion. HA can also be used as bioadhesives
XX for haemostatic sealing and healing of wounds and surgical incisions; and
XX as biomaterials that provide sustained delivery of encapsulated drugs, to
XX wounds, ulcers, injuries or surgical sites. The present sequence can
XX therefore be used to produce HA.
XX
XX Sequence 702 AA;
XX
XX Query Match 60.5%; Score 3079; DB 21; Length 702;
XX Best Local Similarity 82.4%; Pred. No. 1.2e-227;
XX
XX

Matches 579; Conservative 58; Mismatches 58; Indels 8; Gaps 3:
QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFOIIKCKEL-----STNS---YVS 53
DB 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVEFOIIKCKELSAHPSVNSAHLVPN 60
QY 54 EDKNSVCDSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKKSNAERKVELVP 113
DB 61 KEEKVNVCDSPDIATQALLSNVKKLVSDSEKNTLKNKWLTERKSENAEYRAVALVP 120
QY 114 KDFPKDLVLAPDPDVNDFTWYKNRKKSLGKIPVKNKNGISLIIPFNRSRIIDITLACL 173
DB 121 KDFPKDLVLAPDPDVNDFTWYKNRKKSLGKIPVKNKNGISLIIPFNRSRIIDITLACL 180
QY 174 VNOKTNPFEVYVADGSGKENLTIQVQKQKLDIYVRQKQYQYQYQYQYQYQYQYQY 233
DB 181 VNOKTNPFEVYVADGSGQEDLSPIIRQYENKLDIYVRQKQYQYQYQYQYQYQYQY 240
QY 234 DFVSIILDCDAPQOLWVHSYLTLELNDIVLIGPRKYVDTHNITABOFLNDPVLIESLP 293
DB 241 DFVSIILDCDAPNPLWVHSYVAELLEDDLTIIIGPRKYIDTQHDIPKDFLNNASLLES 300
QY 294 ETATNNPSITSKGNISLDWRLHFKKTDNLRLCDSFPRYFVAGNVAFSKWLKNKVGWFD 353
DB 301 EVKTNNSVAAGGEGTVDLWRLHFKKTDNLRLCDSFPRYFVAGNVAFSKWLKNKVGWFD 360
QY 354 EEFNHGGEDVEFGYRLFAGCGFRRVDDGMAIHQEPGPKENETEREAGKSITLKIYKEK 413
DB 361 EEFNHGGEDVEFGYRLFAGCGFRRVDDGMAIHQEPGPKENETEREAGKSITLKIYKEK 420
QY 414 VPYIYRKLLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCICNDGST 473
DB 421 VPYIYRKLLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALNQTVVDLEVCICNDGST 479
QY 474 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAVSFAGYIIGOLDSDDDYLEDAVELC 533
DB 480 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAVSFAGYIIGOLDSDDDYLEDAVELC 539
QY 534 LKEFLKDKTLACVYTTNRVNPDPGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 593
DB 540 LKEFLKDKTLACVYTTNRVNPDPGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 599
QY 594 TDGPNENIENAVDYDMFLKLVSEVGKFKHLNKICYNRVLHGDNSTSKKLGIOKKNHFVVVN 653
DB 600 TDGPNENIENAVDYDMFLKLVSEVGKFKHLNKICYNRVLHGDNSTSKKLGIOKKNHFVVVN 659
QY 654 QSLNRQGINNYNYDFDLDLDESRYIFNKTAEYQEMDMKDL 696
DB 660 QSLNRQGINNYNYDFDLDLDESRYIFNKTAEYQEMDMKDL 702
RESULT 7
ABB48565
ID ABB48565 standard; Protein; 327 AA.
XX
AC ABB48565;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1269.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000PR-0004629.

```
XX PA (INSP ) INST PASTEUR.
XX PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
XX PI Dussurget O, Chetoui L, Medjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Kref J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;
XX DR WPI; 2002-010914/01.
XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX PT and prevention of Listeria and related bacterial infections, and
XX PT related polypeptides -
XX PS Claim 6; SEQ ID No 1270; 192pp; French.
XX CC The present invention relates to the genome sequence of Listeria
XX CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX CC it are useful for selecting probes and primers for detecting genes in L.
XX CC monocytogenes and related organisms, and for studying genetic
XX CC polymorphisms and other genomes. The present sequence is a protein
XX CC encoded by the genome sequence of the present invention. Proteins
XX CC expressed from the genome sequence are useful for raising specific
XX CC antibodies, identification of L. monocytogenes and related organisms, and
XX CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX CC B12. The genome sequence and proteins encoded by it are also useful for
XX CC selecting compounds that regulate gene expression and cell replication
XX CC and modulate L. monocytogenes-related diseases. In addition, the genome
XX CC sequence and proteins encoded by it are useful in pharmaceutical and
XX CC vaccines compositions for the treatment or prevention of infections by L.
XX CC monocytogenes and related organisms.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 327 AA;
Query Match 5.7%; Score 292.5; DB 23; Length 327;
Best Local Similarity 27.5%; Pred. No. 5.5e-14;
Matches 98; Conservative 44; Mismatches 133; Indels 81; Gaps 9;
Qy 433 PLVSIYIPAYCANYIQCVDYSDALNQTVDLEVCICNDGSDTDLVINKLYGNPNRVRI 492
Db 3 PLVSVIIPVYVNEYKVRKCLDSVLEQYVHNLEVIIVNDGATDNSAKVIKTSIDN--RIRY 60
Qy 493 MSKPNGGIASNAVSPFAGYIYQDSDDDYLEPDAVELCLKEFLKDKTLACVYTTNRN 552
Db 61 FEKENAGQATARNGLDVATGDYIVMVDSDYISKNLVETCL-DTVQKTNADLVLTFSYN 119
Qy 553 VNPQGS-----LIANGYNWPEFSREKLTMTAMIAHHFRFTTRAHNLTDGP--- 597
Db 120 VNQEGKQYIKRDKGIKVLDAQPTPNKFKYQADLWKG-----SRFPVGYWYEDLGIIPV 173
Qy 598 -----NENTENADVY---DMFLKLSVGFKFLKLCYKNRVLHGDNTSIKKLGIQKN 647
Db 174 VTLKADNPVKIQDALYIYITRADSQSNIQQVDFLDVVI--MLENVETELKLGII----- 227
Qy 648 HFVVVNSLNRQGINNYNDKFDLDESRYIFNKTAPEQEMDKOL---KLIONKDA 704
Db 228 -----YEESKOQLAYLYIEHLYRLVL 249
Qy 705 KIAYSIFPNTLGLVKKLNINIEYKNIFVILIIVHDKNHILTPDIKKEILAFYHKH 760
Db 250 RKAIYITNKQKKKLIKIKISTIQEKFPMWGSYPYQAGGKLTATLKKKALWYLHH 305
RESULT 8
AAU33454
ID AAU33454 standard; Protein; 706 AA.
```

```
XX AC AAU33454;
XX DT 14-FEB-2002 (first entry)
XX DE Enterococcus faecalis cellular proliferation protein #90.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Enterococcus faecalis.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX N-PSDB; AAS51313.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 4950; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 706 AA;
Query Match 5.4%; Score 276.5; DB 22; Length 706;
Best Local Similarity 21.5%; Pred. No. 3.1e-12;
Matches 151; Conservative 115; Mismatches 268; Indels 167; Gaps 25;
Qy 260 DNDIVILGPKRYVD--THNITAEQFLNDPYLIESLPETATNNPNPITSKGNISLDMR--- 314
Db 2 NEDIKVIQFSDIYRDKATNNLTITGWALDTITKES--PTFTINNQVSAYNRIQVRLREDV 59
Qy 315 -----LEHFKKTDNLRLCDSPPFFYFVAGNVAFSKEWLNK-----VGWF 352
Db 60 NQIYOTEPAIEAGFVVTLEGIKOKKVL-----PPHFQSSAHVVTVDFFLNKKYVIPGTE 114
```

QY	353	DEEFNHGGEDVEFGYRLFAKGCFFRVIDGGMAIHQEPGKNETEREAGKSITLIKIVKE	412
Db	115	DKVTRLN-----IRAKGFKYMAKNGIS-HTIQRAKIEKLRNQA--SYLNLWLARN	161
QY	413	KVPYIYRKLLPTEDSHIHRIPLVSIYPAYNC-ANYIQRVDSALNQTVDLEVCICNDG	471
Db	162	EVLDI--EAMTQEIATFHQPKISAMPVYVVEEKWLRCLCIDSILNQVYTNMELCMADDA	219
QY	472	STD-NTLEVINKLYGNPRVRIM-SKPNGGIASASNAVSAFAGYYIGOLDSDDYLEPDA	529
Db	220	STDPNVKKILTEYOQLDERIRVVRONGHISEATNSALAIATGEFVALLDNDDELAINA	279
QY	530	VELCLKEFLKDKTLACVYTTNRNPDGSLIANGYNNWPEFSREKLTAMIAHFRMTTIR	589
Db	280	FYEVVKVLNENPELDLIYSDEKIDMDGNRSDPAFK-PDWSPLLGLTNYISHLGVYRRS	338
QY	590	AHWLTDFGNENIENAVDYDMFLKSE---VGKFKHLNKCICYNRVLHGDNTSIRK--LGIO	644
Db	339	ILEEIGGFRKGYEGSQDYDLVLFTEKTKERTHPIKVLVYWRMLPTSTAVDQSGGYA	398
QY	645	KNHFVVVQSLNRQGINYN-----YDKFDDLDESRYIFNKTAEQEEMDMKDL	696
Db	399	FEAGLRAVDALVRGINGHATHGAANGLYDVIYDI-ESEK-----	438
QY	697	KLIQNKDAKTAVSIFYPNTLNG---LVKLNLIIEYKNKIFVILLHVDKNHLPDIKKEI	753
Db	439	-----LVSIIP-TKNGYKDVQRCVSSIIIEKTYQNYELIMADNSTDPKMHELY	487
QY	754	LAFYKHQVNLNN-DISYYTNSRLIKTEAHLNKLSQLNLCNCEYIIFDNDHSLFVK	812
Db	488	AKFEQOLPGRFVESIDIPF-----NFTINNRKAHKAHGEVLLFLNNDTEVIT	536
QY	813	ND-----SYA-----	821
Db	537	ENWLTLMVSAFQERIGCVGAKLLYPNNVQAHGVLGLGVAGHGYGPHGLDGFGR	596
QY	822	YDVGNFSAHLDWT-----EKINAHPPPKLIKTYFNDNL	858
Db	597	LAINVYSAVTAACLLMKKADFVAGGFEAFVAFNDVDL	637
RESULT 9			
ID	AAU34918	standard; Protein; 715 AA.	
XX	AAU34918;		
AC	AAU34918;		
DT	14-FEB-2002	(first entry)	
XX	Enterococcus faecalis	cellular proliferation protein #205.	
DE	Antisense; prokaryotic	cellular proliferation protein;	
XX	antibiotic; antibacterial;	drug design.	
KW	Enterococcus faecalis.		
XX	WO200170955-A2.		
OS	27-SEP-2001.		
PN	21-MAR-2001;	2001WO-US09180.	
XX	21-MAR-2000;	2000US-191078P.	
PR	23-MAY-2000;	2000US-206848P.	
PR	26-MAY-2000;	2000US-207727P.	
PR	23-OCT-2000;	2000US-242578P.	
PR	27-NOV-2000;	2000US-253625P.	
PR	22-DEC-2000;	2000US-257931P.	
PR	16-FEB-2001;	2001US-269308P.	
XX	(ELIT-) ELITRA PHARM INC.		
PA	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;		
PI			
PI	Yamamoto RT, Xu HH;		
XX	WPI; 2001-611495/70.		
DR	N-PSDB; AAS52777.		
XX	New polynucleotides for the identification and development of		
PT	antibiotics, comprise sequences of antisense nucleic acids -		
XX	Example 3; Seq ID No 10511; 511pp; English.		
PS	The invention relates to antisense inhibitors of genes essential to		
CC	prokaryotic cellular proliferation, their use in identifying the		
CC	genes, their use in the discovery of novel antibiotics, the essential		
CC	genes themselves and the encoded proteins. The prokaryotes used are		
CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella		
CC	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The		
CC	invention is also useful for the identification of potential new targets		
CC	for antibiotic development. The antisense nucleic acids can also be used		
CC	to identify proteins used in proliferation, to express these proteins,		
CC	and to obtain antibodies capable of binding to the expressed proteins.		
CC	The proteins can be used to screen compounds in rational drug discovery		
CC	programmes. The antisense nucleic acid sequence is also useful to screen		
CC	for homologous nucleic acids which are required for cell proliferation in		
CC	a wide variety of organisms. The present sequence represents an		
CC	essential prokaryotic cellular proliferation protein.		
CC	Note: The sequence data for this patent did not form part		
CC	of the printed specification, but was obtained in electronic		
CC	format directly from WIPO at		
CC	ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 715 AA;		
SQ	Query Match 5.4%; Score 276.5; DB 22; Length 715;		
	Best Local Similarity 21.5%; Pred. No. 3.2e-12;		
	Matches 151; Conservative 115; Mismatches 288; Indels 167; Gaps 25;		
QY	260	DNDLVLIGPRKYVD--TNITAEOLFNDPILYIESPETATNNPSITSKGNISLDWR---	314
Db	4	NEDIKVFDSIYRKATNNLTITGWALDTIKES--PTFTINNENQVSAYNIQRLREDV	61
QY	315	-----LEHFKKTDNLRCLDSPFRFYFVAGNVAFSKEWLNK-----VGWF	352
Db	62	NOIYOTEPAIEAGFVVTLEGIKQKKVL-----PFHQSSAHVTVVDFPLNKKYVIPGTE	116
QY	353	DEEFNHGGEDVEFGYRLFAKGCFFRVIDGGMAIHQEPGKNETEREAGKSITLIKIVKE	412
Db	117	DKVTRLN-----IRAKGFKYMAKNGIS-HTIQRAKIEKLRNQA--SYLNLWLARN	163
QY	413	KVPYIYRKLLPTEDSHIHRIPLVSIYPAYNC-ANYIQRVDSALNQTVDLEVCICNDG	471
Db	164	EVLDI--EAMTQEIATFHQPKISAMPVYVVEEKWLRCLCIDSILNQVYTNMELCMADDA	221
QY	472	STD-NTLEVINKLYGNPRVRIM-SKPNGGIASASNAVSAFAGYYIGOLDSDDYLEPDA	529
Db	222	STDPNVKKILTEYOQLDERIRVVRONGHISEATNSALAIATGEFVALLDNDDELAINA	281
QY	530	VELCLKEFLKDKTLACVYTTNRNPDGSLIANGYNNWPEFSREKLTAMIAHFRMTTIR	589
Db	282	FYEVVKVLNENPELDLIYSDEKIDMDGNRSDPAFK-PDWSPLLGLTNYISHLGVYRRS	340
QY	590	AHWLTDFGNENIENAVDYDMFLKSE---VGKFKHLNKCICYNRVLHGDNTSIRK--LGIO	644
Db	341	ILEEIGGFRKGYEGSQDYDLVLFTEKTKERTHPIKVLVYWRMLPTSTAVDQSGGYA	400
QY	645	KNHFVVVQSLNRQGINYN-----YDKFDDLDESRYIFNKTAEQEEMDMKDL	696
Db	401	FEAGLRAVDALVRGINGHATHGAANGLYDVIYDI-ESEK-----	440
QY	697	KLIQNKDAKTAVSIFYPNTLNG---LVKLNLIIEYKNKIFVILLHVDKNHLPDIKKEI	753
Db	441	-----LVSIIP-TKNGYKDVQRCVSSIIIEKTYQNYELIMADNSTDPKMHELY	489
QY	754	LAFYKHQVNLNN-DISYYTNSRLIKTEAHLNKLSQLNLCNCEYIIFDNDHSLFVK	812


```
Db 490 AKFEQQLPGRFVESIDIPF-----NFSTINNRAAKAHGEVLLFLNNDTEVIT 538
Qy 813 ND-----SYA-----YMKK 821
Db 539 ENWJLTMVSAQQBRIGCVGAKLPLPNNVVOHAGVILGLGVAGHGHYGYPHGDLYGFR 598
Qy 822 YDVGMFSAETHDWI-----EKINAHPPFKKLKITYFNDNL 858
Db 599 LAINVNSAVTAACLLMKKADFDAVGGEFAFTVAENDVDL 639

RESULT 10
ID ABP26805 standard; Protein; 321 AA.
XX
AC ABP26805;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 2786.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN W0200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
WPI; 2002-352536/38.
DR N-PSDB; ABN67436.
XX
New Streptococcus protein for the treatment or prevention of infection
or disease caused by Streptococcus bacteria, such as meningitis, and
for detecting a compound that binds to the protein -
XX
Claim 1; Page 3429; 4525pp; English.
XX
The invention relates to a protein (ABP25413-ABP30895) from group B
streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
the specification. The proteins have antibacterial and antiinflammatory
activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
antibodies that bind (I) are used in the manufacture of medicaments for
the treatment or prevention of infection or disease caused by
Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
Nucleic acids encoding (I) are used to detect Streptococcus in a
biological sample. (I) is used to determine whether a compound binds to
(I). A composition comprising (I) or a nucleic acid encoding (I), may be
used as a vaccine or diagnostic composition. The disease caused by
Streptococcus that is prevented or treated may be meningitis. Nucleic
acid encoding (I) may be used to recombinantly produce (I) and may be
used in gene therapy. Antibodies to (I) are used for affinity
chromatography, immunoassays, and distinguishing/identifying
Streptococcus proteins.
XX
Sequence 321 AA;
```

```
Query Match 5.0%; Score 253; DB 23; Length 321;
Best Local Similarity 25.1%; Pred. No. 5.8e-11;
Matches 94; Conservative 64; Mismatches 142; Indels 74; Gaps 15;
Qy 435 VSIYIPAYCANYTORCVDSALNQTVVDLEVCICNDGSTDNTLEVINKLYGNPNRVRIMS 494
Db 4 VSIIPVYNQSFLENCIESVLAQTSYNLEILVNDGSDNSGDICDYSEIDGRIFVFH 63
Qy 495 KPNGGIASANAASFAKGYIGOLDSDYL-EPDAVELCL--KEFLKDKTLACVYTTN 550
Db 64 KNGGLSDARNYGISRATGDIYLLDSDYLYKEDAIERMVEFSEKYEIVLGC-YVEK 122
Qy 551 R-----NVNPDGSLIANGYNPFESREKLTAMTAIAHFRMFTIRAWHLTDGFENENAV 605
Db 123 REQHIINIVLEDEMI-----ETISPV-----QAQNIY 150
Qy 606 DYDMFLKSEVGKFKHLNKKICYNRVLHGDNTSIIKLGLOKKNHFVVVQSLNRQGINY 665
Db 151 NYDAYRAIFTVAH-----NKL-YKRELF--STLCYPVGLKHEDEFTYKLYLAKAKNIIFR 203
Qy 666 YDKFDDLDSESKYIFNK----TAEQEEMDKLKLKIONKDAKIAVSIFYPNTL---N 717
Db 204 YNTY-----AYRIRENSINTGSYN-----IKRLHVAEALKERYLLEKYPDLVFQSER 251
Qy 718 GLVKKLN-NIIEYNKNIEVIIHVDKNHLTPDIKKEILAFYHKHQNILLNNDISYTSN 776
Db 252 ALIKTMEVNLIELYKNFYKEFHT---LKTEYKKTIFDFIKKQRMLLKIKYLLAYCVVH 307
Qy 777 RLKTEAHLNINK 790
Db 308 FKILNCKRKKKINK 321

RESULT 11
AAB96313
ID AAB96313 standard; Protein; 298 AA.
XX
AC AAB96313;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative glycosyltransferase, involved in cell wall biogenesis #1.
XX Hyperthermophilic archaeon; hyperthermophilic protein.
XX Pyrococcus abyssi.
XX OS Pyrococcus abyssi.
XX PN FR2792651-A1.
XX PD 27-OCT-2000.
XX PF 21-APR-1999; 99FR-0005034.
XX PR 21-APR-1999; 99FR-0005034.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX Querellou J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
proteins useful in industry -
XX Claim 7; Pages 981-982; 1657pp; French.
XX The present invention relates to the genomic sequence of Pyrococcus
abyssi (see AAF86431 and AAF41223-7) and P. abyssi proteins. P. abyssi is
a hyperthermophilic archaeon, which is isolated from deep-sea
hydrothermal vents. The present sequence is one such P. abyssi protein.
XX The proteins of the present invention have various potential industrial
```

CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.

XX Sequence 298 AA;
SQ Query Match 4.9%; Score 251; DB 22; Length 298;
Best Local Similarity 24.3%; Pred. No. 7.3e-11;
Matches 81; Conservative 66; Mismatches 136; Indels 50; Gaps 8;

QY 433 PLVSIIPAYNCANYIQCVDLSALNQIVVDLEVCICNDGSDTNTLEVINKLYGNPRVR- 491
DB 4 PIVSVIPTYNANLRLRAIASVNLQKFDLIVVDASTDNTPEVESI--EDGRIRY 61
QY 492 IMSKPNGGIASNAASFAKGYIGOLDSDDYLEDPDAVELCLKEFLK-DKTLACVYTN 550
DB 62 IRLKNSGGPIARNIGIKKAKGRFIALDDDDWLPRLHVQVRKFNELGKEFGVYGGF 121
QY 551 RNVNPDGSLIANGYNMPEFSREKLTAMTAHFF-----RMFTTRAWHLTDGFNENIENAV 605
DB 122 YVVSQDGRIL--GKRLPK-HRGDIYSHLLKENFISPTLLIRRECFKAGLFDPLRSSQ 178
QY 606 DYDMFLKLSEVGKFKHLNKICYNRVLHGDNWTSIKLGIQKKNHFVVVQNSLNRQGINYN 665
DB 179 DWMMLRIARYKFDVYDEIIAKYVHGKQISF----- 211
QY 666 YDKFDDLDSEKRYIFNKTAEYQEMDMKDLKLIQNKDAKIAYSFVPNTLNLGLVKLN- 724
DB 212 -----NMKKYIPGRERLIRKHLDIWKNPKILSIHLSQGLLLLLSNNTGKGLKYLTY 263
QY 725 --NTEIENKIFVILHVDKNHLPDIIKKEILA 755
DB 264 SIATAPLENYIMILLKALDSRTVEYIKRILS 296

RESULT 12
AAI68974
ID AAY68974 standard; Protein; 322 AA.
XX AC AAY68974;
XX DT 30-MAY-2000 (first entry)
XX DE CpsII protein which has glycosyltransferase activity.
XX KW Capsular gene cluster; serotype 1; polysaccharide biosynthesis;
KW capsular component; antigen; regulation; chain length determination;
KW complement-mediated opsonophagocytosis; serotype-specific detection;
KW antigen; vaccine; Streptococcal disease; CpsII; CpsI; CpsII;
KW CpsII; CpsII; CpsII; CpsII; glycosyltransferase; CP polymerase.
XX OS Streptococcus suis.
XX PN WO200005378-A2.
XX PD 03-FEB-2000.
XX PF 19-JUL-1999; 99WO-NL00460.
XX PR 22-JUL-1998; 98EP-0202465.
XX PR 22-JUL-1998; 98EP-0202467.
XX PA (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.
XX PI Smith HE;
XX WPI; 2000-195104/17.
XX N-PSDB; AAZ60930.
XX PT New nucleic acid containing the capsular gene cluster of Streptococcus
PT suis, used for serotype-specific detection and to generate antigens or

PT mutants for vaccination -
XX Disclosure; Fig 4; 144pp; English.

CC The proteins AAY68970-76 are encoded by the capsular gene cluster of
CC Streptococcus suis serotype 1. The genes in this cluster are involved
CC in polysaccharide biosynthesis of capsular components and antigens. The
CC proteins have glycosyltransferase activities (CpsII, CpsI, CpsII,
CC CpsII, CpsII) and CP polymerase activities (CpsII). The capsule confers
CC bacterium resistance to complement-mediated opsonophagocytosis. The
CC gene cluster is used as a source of probes and primers for
CC serotype-specific detection of S. suis and is also useful for
CC recombinant production of the proteins. The proteins are then useful
CC for producing antigens that can be used in vaccines, for controlling
CC or eradicating a Streptococcal disease, in humans or animals,
CC e.g. against S. suis in pigs.

XX Sequence 322 AA;
SQ Query Match 4.8%; Score 246; DB 21; Length 322;
Best Local Similarity 25.3%; Pred. No. 2e-10;
Matches 92; Conservative 58; Mismatches 120; Indels 94; Gaps 15;

QY 434 LVSIIIPAYNCANYIQCVDLSALNQIVVDLEVCICNDGSDTNTLEVINKLYGNPRVRIM 493
DB 4 LISVIVPIYNVQDYLDKICINSIINQTYNLEVILVNDGSDTNTLEVINKLYGNPRVRIM 63
QY 494 SKPNNGGIASNAASFAKGYIGOLDSDDYLEDPDAVELCLKEFLKDKTLACVYTNRV 553
DB 64 KKNGLADARNGLHATGKYAFVDSDDY-----TEVAMFERMHD-----NI 107
QY 554 NPDSGLIA-----NGY-----NWPEFSREKLTAMTAHFFMFTIRAWHLTDGF 597
DB 108 TEYNADIAEIDFCLVDENGTYTKKRNNSFHVLTRETVKEFLS----- 150
QY 598 NENIENAVDYDMFLKLSEVGKFKHLNKICYNRVLHGDNWTSIKLGIQKKNH-- 648
DB 151 GSNIEI-----NVCKLYSRDIKDIKFIQNNRSIGEDLLFNVLNNVTRVVDREYY 206
QY 649 -FVVVQNSLNRQGINYNKDFDLDSEKRYIFNKTAEYQEMDMKDLKLIQNKDAKIA 707
DB 207 NVVIRNSSLLINQKFSINNIDLVRLE---NYPFKLKREFSHYF---DAKVIKEK----- 254
QY 708 VSIFPNTLNLGLVKLNIIENY--KNIEFVILHVDKNHLPDIIKKEILA 752
DB 255 -----VKCLNKMVSTDCDLDNEFLFLESYKRIYRPFKAKRILS--RKHLV 300
QY 763 NILL 766
DB 301 TLXL 304

RESULT 13
AAI14078
ID AAW14078 standard; Protein; 324 AA.
XX AC AAW14078;
XX DT 25-SEP-1997 (first entry)
XX DE S.thermophilus exopolysaccharide biosynthesis enzyme EpsH.
XX KW Exopolysaccharide biosynthesis; eps operon; lactic acid bacterium;
KW epsR; epsR; epsR; epsR; epsR; epsR; epsR; epsR; epsR; epsR;
KW epsR; epsR; epsR; epsR; epsR; epsR; epsR; epsR; epsR; epsR;
XX OS Streptococcus thermophilus strain Sfi6.
XX PN EP750042-A1.
XX PD 27-DEC-1996.
XX PT 20-JUN-1995; 95EP-0201669.

XX	AA54079;	
AC		
XX		
DT	27-MAR-2000	(first entry)
XX		
DE	Enzyme EPS5 involved in exopolysaccharide biosynthesis.	
XX		
KW	Exopolysaccharide; EPS; ESP enzyme; EPS1; EPS2; EPS3; EPS4; EPS5; EPS6;	
KW	EPS7; EPS8; EPS9; EPS10; EPS11; Lactobacillus helveticus strain LH59;	
KW	activated D-galactose pyranose; saccharide; beta-glycosyltransferase;	
KW	undecaprenyl-phosphate-glycosyl-1-phosphate-transferase;	
KW	alpha-glycosyltransferase; EPS polymerase; glycosyltransferase;	
KW	phosphofuranose; transporter; food; fermented milk product; yoghurt;	
KW	cheese; flavour stability; organoleptic property.	
XX		
OS	Lactobacillus helveticus.	
PN		
XX	WO9962316-A2.	
XX		
PD	09-DEC-1999.	
XX		
PF	22-APR-1999; 99WO-EP02841.	
XX		
PR	22-APR-1998; 98EP-0201310.	
PR	22-APR-1998; 98EP-0201311.	
PR	22-APR-1998; 98EP-0201312.	
XX		
PA	(NEST) SOC PROD NESTLE SA.	
XX		
PI	Stingele F, Germond JE, Lamothe G;	
XX		
DR	WPI; 2000-097267/08.	
DR	N-PSDB; AA245259, AA54082, AA54083, AA54084; AA54085.	
XX		
PT	New recombinant enzymes for synthesis of exopolysaccharides,	
PT	particularly in lactic acid bacteria, for improving properties of	
PT	fermented milk products	
XX		
FS	Claim 9; Page 122-123; 162pp; French.	
XX		

Search completed: January 4, 2003, 02:23:03
Job time : 123 secs

Db	63	IIDKNEGYSIARNTALKNVNSKFVTFVDSDDFVDSNLFVLYSGFKDAKVDLTISGR	122
QY	539	K-----DKTLACVYTTN--RVNPDGSLIAN-----GYNWPEFSREKLTAMIAHFRMET	587
Db	123	KLNTNNKLEEKFSNVTRSAEDVISLLNETGPOGYLWKNKMF-----TSIIKKYSLLLD	178
QY	588	IRAWHLTDGFNENIENAVDYMFLKLVSEVGKFKHLNKICYNRVLHG-----DNTSIKKL	641
Db	179	PKTF-----MAEDLLFCIQYLKESKKVRSNYCDYVQDSNSMNGLSPTKNNRRYKKI	233
QY	642	GIOKKHNVVVVNSLRQGINYYNDKDDLSRKY---IFNKTAEQEEMDLKDLKL	698
Db	234	---FNNVFFALNKILN-----EIKNNPKKYSIAIINVKARLAREYSIFLRLLL	278
QY	699	IQN-KDAKIAVSIFYPNTLNLGLVKLNIIETYNKNIFVIIHLVDKNHLLTPDIKKEILA--	755
Db	279	LNNDKNNKKLOKEVWQ-----AYKLRYVFKSKI-----FVSKKMFFACT	319
QY	756	FYHKHQVNI	764
Db	320	IYCPHLVHI	328

XX	AAV54079;	
AC		
XX		
DT	27-MAR-2000	(first entry)
XX		
DE	Enzyme EPS5 involved in exopolysaccharide biosynthesis.	
XX		
KW	Exopolysaccharide; EPS; ESP enzyme; EPS1; EPS2; EPS3; EPS4; EPS5; EPS6;	
KW	EPS7; EPS8; EPS9; EPS10; EPS11; Lactobacillus helveticus strain LH59;	
KW	activated D-galactose pyranose; saccharide; beta-glycosyltransferase;	
KW	undecaprenyl-phosphate-glycosyl-1-phosphate-transferase;	
KW	alpha-glycosyltransferase; EPS polymerase; glycosyltransferase;	
KW	phosphofuranose; transporter; food; fermented milk product; yoghurt;	
KW	cheese; flavour stability; organoleptic property.	
XX		
OS	Lactobacillus helveticus.	
XX		
PN	WO9962316-A2.	
XX		
PD	09-DEC-1999.	
XX		
PF	22-APR-1999; 99WO-EP02841.	
XX		
PR	22-APR-1998; 98EP-0201310.	
PR	22-APR-1998; 98EP-0201311.	
PR	22-APR-1998; 98EP-0201312.	
XX		
PA	(NEST) SOC PROD NESTLE SA.	
XX		
PI	Stingele F, Germond JE, Lamothe G;	
XX		
DR	WPI: 2000-097267/08.	
DR	N-PSDB; AA245259, AA54082, AA54083, AA54084; AA54085.	
XX		
PT	New recombinant enzymes for synthesis of exopolysaccharides,	
PT	particularly in lactic acid bacteria, for improving properties of	
PT	fermented milk products	
XX		
FS	Claim 9; Page 122-123; 162pp; French.	
XX		
CC	AA54075-85 represent enzymes involved in the biosynthesis of	
CC	exopolysaccharides (EPS). These enzymes are designated EPS1-EPS11. and	
CC	are encoded by open reading frames eps1-eps11. The enzymes are isolated	
CC	from Lactobacillus helveticus strain LH59. The proteins are used	
CC	in a method for the synthesis of EPS, which includes at least one step	
CC	of forming a bond (alpha or beta-isomer) between C-1 (carrying the	
CC	reducing aldehyde function, of an activated D-galactose pyranose), and	
CC	a phosphate on a lipophilic or proteinaceous primer. Stepwise synthesis	
CC	of EPS occurs with, in each step, addition of a new sugar unit, through	
CC	its hemi-acetyl function, to an alcoholic hydroxyl of a second sugar	
CC	unit, present at the end of a chain of sugar residues bonded to the	
CC	primer. EPS1 has homology with an undecaprenyl-phosphate-glycosyl-1-	
CC	phosphate-transferase; EPS2, EPS3 and EPS4 are alpha-	
CC	glycosyltransferases; EPS5 and EPS6 have homology with a beta-	
CC	glycosyltransferase; EPS7 has homology with an EPS polymerase; EPS8	
CC	is a glycosyltransferase; EPS9 catalyses the transfer of phosphofuranose	
CC	onto the following repetitive unit; EPS10 transports the polysaccharides	
CC	formed; and EPS11 is involved in the synthesis and export of formed	
CC	polysaccharides. The EPS enzyme are used to improve properties of foods,	
CC	particularly fermented milk products such as yoghurt and cheese,	
CC	e.g. their organoleptic properties and flavour stability.	
XX		
SQ	Sequence 339 AA;	
	Query Match	4.5%; Score 230.5; DB 21; Length 339;
	Best Local Similarity	25.2%; Pred. No. 3.4e-09;
	Matches 93; Conservative	64; Mismatches 131; Indels 81; Gaps 17;
QY	433	PLVSIYIPAYNCANYIQRCVDSALNQTVVD-LEVICNDGSDTNTLEVINKLGNPRVR 491
Db	4	PLISVIVPAYNAEKSIKCFISLLRONIKESYKIIIVNDGSDTDTTEI-KSFDNPKIS 62
QY	492	IMSKPNGGIASNAASVFAKGYIIGQLDSDDYLE-----PDA--VELCLKEFL 538

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:22:39 ; Search time 66 seconds
(without alignments)
430.199 Million cell updates/sec

Title: US-09-842-484A-2

Perfect score: 5089

Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SARKGENIPVKNFIINSITL 965

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5089	100.0	965	4	US-09-437-277-3
2	3079	60.5	702	4	US-09-437-277-1
3	238.5	4.7	324	1	US-08-597-236-10
4	238.5	4.7	324	1	US-08-746-682A-10
5	227	4.5	727	4	US-09-134-001C-4067
6	223.5	4.4	674	4	US-08-961-083-200
7	210	4.1	348	1	US-08-312-387B-3
8	210	4.1	348	1	US-08-312-387B-11
9	210	4.1	348	1	US-08-683-426-3
10	210	4.1	348	1	US-08-683-426-11
11	210	4.1	348	1	US-08-683-458-3
12	210	4.1	348	1	US-08-683-458-11
13	210	4.1	348	2	US-08-878-360-3
14	210	4.1	348	2	US-08-878-360-11
15	210	4.1	348	3	US-08-478-140B-3
16	210	4.1	348	3	US-08-478-140B-8
17	210	4.1	348	4	US-09-333-412-3
18	210	4.1	348	4	US-09-333-412-11
19	210	4.1	348	4	US-09-338-943-3
20	210	4.1	348	4	US-09-338-943-8
21	209	4.1	337	1	US-08-312-387B-5
22	209	4.1	337	1	US-08-312-387B-12
23	209	4.1	337	1	US-08-683-426-5
24	209	4.1	337	1	US-08-683-426-12
25	209	4.1	337	1	US-08-683-458-5
26	209	4.1	337	1	US-08-683-458-12
27	209	4.1	337	2	US-08-878-360-5

28 209 4.1 337 2 US-08-878-360-12 Sequence 12, Appl
29 209 4.1 337 3 US-08-478-140B-5 Sequence 5, Appl
30 209 4.1 337 4 US-09-333-412-5 Sequence 5, Appl
31 209 4.1 337 4 US-09-333-412-12 Sequence 12, Appl
32 209 4.1 337 4 US-09-338-943-5 Sequence 5, Appl
33 178.5 3.5 990 2 US-08-392-625-20 Sequence 20, Appl
34 178.5 3.5 990 2 US-08-466-961A-20 Sequence 20, Appl
35 176.5 3.5 990 2 US-08-645-193B-15 Sequence 15, Appl
36 174.5 3.4 2710 1 US-08-480-604A-6 Sequence 6, Appl
37 174.5 3.4 2710 2 US-08-405-496A-6 Sequence 6, Appl
38 174.5 3.4 2710 4 US-08-915-136-6 Sequence 6, Appl
39 174.5 3.4 2710 4 US-08-957-310-6 Sequence 6, Appl
40 173.5 3.4 281 4 US-08-961-083-196 Sequence 196, App
41 167.5 3.3 270 4 US-08-961-083-198 Sequence 198, App
42 156.5 3.1 633 2 US-08-648-298-2 Sequence 2, Appl
43 156.5 3.1 10182 4 US-09-134-001C-3159 Sequence 3159, Ap
44 156 3.1 418 4 US-09-134-001C-4051 Sequence 4051, Ap
45 153.5 3.0 956 4 US-09-134-001C-4452 Sequence 4452, Ap

ALIGNMENTS

RESULT 1

US-09-437-277-3
; Sequence 3, Application US/09437277
; Patent No. 644447
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 5820.551
; CURRENT APPLICATION NUMBER: US/09/437,277
; CURRENT FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
; SEQ ID NO 3
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-437-277-3

Query Match 100.0%; Score 5089; DB 4; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFEQIIRCKEKLSTNSYVSEDKKNSV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFEQIIRCKEKLSTNSYVSEDKKNSV 60
Qy 61 CDSSLDIATQLLLSNVKKLTLSSEKSLKNKWSITCKKSENAIRKVELVPKDFPKDL 120
Db 61 CDSSLDIATQLLLSNVKKLTLSSEKSLKNKWSITCKKSENAIRKVELVPKDFPKDL 120
Qy 121 VLAPLDHVNDFTWYKNRKKSLGKIPVKNKIGLSIIIPFNRSRLDITLACLVNQKTNV 180
Db 121 VLAPLDHVNDFTWYKNRKKSLGKIPVKNKIGLSIIIPFNRSRLDITLACLVNQKTNV 180
Qy 181 PFEVVVADGSKENLLTIVQKYEOKLDIKYVRQKDYQOLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQKYEOKLDIKYVRQKDYQOLCAVRNLGLRTAKYDFVSILD 240
Qy 241 CDMAPOQLVWHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLPETATNNN 300
Db 241 CDMAPOQLVWHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLPETATNNN 300
Qy 301 PSITSGKNISLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKEWLNKVGWDFDEFNHWG 360
Db 301 PSITSGKNISLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKEWLNKVGWDFDEFNHWG 360
Qy 361 GEDVEFGYRLFAKCGFFRVIDGGMAIHOEPGKNETREAGKSIITLIVKEKVPYIYRK 420
Db 361 GEDVEFGYRLFAKCGFFRVIDGGMAIHOEPGKNETREAGKSIITLIVKEKVPYIYRK 420

69 IDLDSNGHARNIALEBETVPEMFELDADDELASAIATTFYLEKFNNTDGLIAPHST-128
549 TNRNVPDGLSLANGY-----NWPEFSREKLTAMIAHFRMFTIRAWHLTDGCFNENIEN 603
129 TORPOQVLDLDRVVEYFNAKENINSFLRKQSAACNII---FRTAIRAHHR--FNEULNT 183
604 AVYDMFLKSLSEVGFKHLKIKYNRVLHGDNTSIKGLIQKKNHFVVVYNQSLNRQGINY 663
184 YDWSVLEY-----MKYVNF-----VRI 203
664 YN-----YDKFDDLDSESKRYIFNKTAEOEEMDL-KDLKLIQNDAKIAVSIFY- 712
204 FNPFFFRGEVDPFETLTL-----EQNFDILFKDY-----VNSFYD 241
713 -----PNTLGNLVKLANII-----EYKNKIFVILHVDKNHLPDIPKKEILAFYH 758
242 AKRATNPVREFIVTKMGNKIANEPTRYDIN-----ERYOHTKDLVELSKFLH 293
759 KHOVN-----ILLNNDI-SYITSNRLIKTEAHLNSI-----NKLSQL-----794
294 VHLVKQKLINKLETILLMNNETDKAFVQFRKTLRHVKNIVLRRKNKERSLYDLTDKE 353
795 -NUNCEYIIFDNDHSLFVNDS-----YAYMKY-----822
354 DNVKPKTIIVPESFG-----KNYDSPKIYIYEMQKYPNYRYIWSFKNPKNVVPGSAEKV 410
823 -----DVGMMNFSAETHDWIEKINAHPPF-----KKLIKTYFND-----NDLRSMNVK 864
411 KRNSAEYQAYSEASH-WVS--NARTPLYLNKKENQTYIQTWHTGTPLKRLANDKMKVVRMP 467
865 GAS 867
468 GTT 470
RESULT 6
US-08-961-083-200
; Sequence 200, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 674 amino acids
; TYPE: amino acid

STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-200
Query Match 4.4%; Score 223.5; DB 4; Length 674;
Best Local Similarity 20.1%; Pred.No. 7.8e-11;
Matches 96; Conservative 89; Mismatches 177; Indels 115; Gaps 17;
QY 451 CVDSALNQVWVLEVCICNDGSTDNTLEVINKLYGNPRVRIMSKPNGGIASASNAVSF 510
Db 1 CLDSITQYKNEIIVVNDGSTDASGEICKEFSEMDHRLIYIEQENAGLSAARNTGLNN 60
QY 511 AKGYIIGQLSDSDYLEPDAVELCLK---BFLKDKTKLACVYTTNRN-----552
Db 61 MSGNYVTFVDSDDWIEQDYVETLYKKIIVEYQADIAVGNYSFNESEGMPYFHLGDSYVE 120
QY 553 -VNPDSLSLANGVWPEFSREKLTAMIAHFRMFTIRAWHL-----TDGFNENIE 602
Db 121 KVDVNSIFENLYE-----TOEMKSFALISAWGKLYKARLFEQLRFDIGKLGEDGY-----171
QY 603 NAVDYDMFLKSLSEVGFKHLKICY-NRVLHGD-----NTSIKKL 641
Db 172 --LNQKVYL-LSE--KVIVLNSLYAVIRKSGLSRVWTEKMMHALVDAMSERITILLANM 226
QY 642 GIOKKNHFFVVVNSLRQGINYNDKFDLDESRYIFNKTAEOEEMDLKDLKLIQN 701
Db 227 GYPLEKHLAVYROMLEVSLAN---GQASGLSDTATY---REFEMKORLLNOLSRQEE 277
QY 702 KDAKIASVIFYPNTLGNLVKLANIIEYKNKIFVILHVD-KNHLTPDIKKEILAF-----756
Db 278 SEKKAIVLAANYGYVDQVLTITKISYCHNRSIRFYLHSDFPNWKQLNKRLEKFPDSEI 337
QY 757 ----YHKHQVNILLNNDISYITSNRLIKTEAHLNSIKNLSQNLNLCYIIFDNDHSLFVK 812
Db 338 INCRTVSEQIS-CYKSDISYTVELRYFIADF---VQEDKALYLDCLVYVTKNLDLDFAT 392
QY 813 N-DSY-----AYMKKYDVGMMNFSAETHDWIEKIN 840
Db 393 DLQDYPLAAVDRFGGRAYFGQEIFNAGVLLVNNAFWKKNNTOKLIDVTNEWHDKVD 449
RESULT 7
US-08-312-387B-3
; Sequence 3, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800

```
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-312-387B-3

Query Match      4.18; Score 210; DB 1; Length 348;
Best Local Similarity 26.4%; Pred. No. 4.1e-10;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

Qy 433 PLVSIYIPAYNCANYIORCVDSALNQTVDLEVCICNDGSDTNTLEVLKLYGNPNRVRI 492
      ||||| : ||| : : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3 PLVSLICAYNVKIFYAQSAAVYVQTNWNLILVDDGSDGTGLAIKDFKQKRSRIKI 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 493 MSK-PNGGIASASNAAV-SFAK-----GYIGQLSDSDYLEPDVAVELCLKEFLKDKTLAC 545
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 LAQAQNSGLIPSLNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 546 V--YTTNRNVPDGSLIA-----NGYNWPEFSREKLTMTMI-----AHFRMTTIRAWHLT 594
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 MGAWLEVLSEEDGNRLARHHKHKIKWKPTRHEDIAAFFFFPGNPINHNTMTIMRRS--VI 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 595 DG---FENENIENAVDYDMFLKLSEVKGPKHLNKCINRVLHGDNTSIK-----KLGI 643
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : : ||
Db 181 DGGLRYDTERWAEDYQFWYDVSKGLRLAYYPEALVKYRLHANOVSSKHSVRQHEIAOGI 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 644 QK--KNHFVVVYNQSLNRQGINVYNDKFDLDESCKYIFNKTAEQY-EEMDM 692
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : : ||
Db 241 QKTARNDFL---QSMGFK-----TRFDSLE---YRQTKAAAYELPEKDL 278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-08-312-387B-11
; Sequence 11, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-426-3

Query Match      4.18; Score 210; DB 1; Length 348;
Best Local Similarity 26.4%; Pred. No. 4.1e-10;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

Qy 433 PLVSIYIPAYNCANYIORCVDSALNQTVDLEVCICNDGSDTNTLEVLKLYGNPNRVRI 492
      ||||| : ||| : : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3 PLVSLICAYNVKIFYAQSAAVYVQTNWNLILVDDGSDGTGLAIKDFKQKRSRIKI 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 493 MSK-PNGGIASASNAAV-SFAK-----GYIGQLSDSDYLEPDVAVELCLKEFLKDKTLAC 545
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 LAQAQNSGLIPSLNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 546 V--YTTNRNVPDGSLIA-----NGYNWPEFSREKLTMTMI-----AHFRMTTIRAWHLT 594
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 MGAWLEVLSEEDGNRLARHHKHKIKWKPTRHEDIAAFFFFPGNPINHNTMTIMRRS--VI 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 595 DG---FENENIENAVDYDMFLKLSEVKGPKHLNKCINRVLHGDNTSIK-----KLGI 643
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : : ||
Db 181 DGGLRYDTERWAEDYQFWYDVSKGLRLAYYPEALVKYRLHANOVSSKHSVRQHEIAOGI 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 644 QK--KNHFVVVYNQSLNRQGINVYNDKFDLDESCKYIFNKTAEQY-EEMDM 692
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : : ||
Db 241 QKTARNDFL---QSMGFK-----TRFDSLE---YRQTKAAAYELPEKDL 278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-08-683-426-3
; Sequence 3, Application US/08683426
; Patent No. 5705367
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,426
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-426-3
```

QY 433 PLVSIYIPAYNCANYIQRCVDSALNQTVDVLEVCICNDGSDTNTLEVINKLIGNNPVRI 492
 ||||| : ||| : : ||| : ||| : ||||| || : ||| :
 Db 3 PLVSLICANYEKFQAQSLAAVNTWRNLDLIVDGGSTDTGLIAKDFQKRDSRIKI 62

QY		493	MSX	-PNGGIASASNAAV-SFAK-----GYIIGQLSDSDYLEPAVELCKEFLDKDTLAC	545
	:::	:	:	:	:
	:	:	:	:	:
D6		63	LAQAQNSGLTSLNIGLDELAKSGGGGGEYIARTDADDIASPGWIEKIVGEMEKORSHIA	122	
	:	:	:	:	:
	:	:	:	:	:
QY		546	V--YTNRNRNVDPGSLIA-----NGYNWPFSREKLTTAMI-----AHHRFMTIRAWHLT	594	
	:	:	:	:	:
	:	:	:	:	:
D6		123	MGAWELEVSPEKDGNRLARHHKHGKTWKPKTRHEDIAAFFPGNPINHNTIMRRS--VI	180	
	:	:	:	:	:
	:	:	:	:	:
QY		595	DG---FNENTENADVDMFLKSEVGKFHLNKICYNRVLGHGDNTSIK-----KLGI	643	
	:	:	:	:	:
	:	:	:	:	:
D6		181	DGGLRYDTERDAEDYQFWDVYSKLGRLAAYPEALVKYRHLANQVSSKHSVRQHEIAOGI	240	
	:	:	:	:	:
	:	:	:	:	:
QY		644	QR--KNHFVVVNOSLRNQGINNYNDOKFDDDLDSRKYIFNKTAEQV-EEMDM	692	
	:	:	:	:	:
	:	:	:	:	:
D6		241	OKYARNDFL---QSMGFK-----TRFDSLE-----YROTKAAPVLPPEKDL	278	
	:	:	:	:	:
	:	:	:	:	:

RESULT 11
US-08-683-458-3

Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

```

:
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0.
:
: Version #1.25
:

```

```

:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/683,458
:

```

APPLICATION NUMBER: 85767-1007
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 20,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800

```

; INFORMATION FOR SEQ ID NO:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 348 amino acids

```

; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-683-458-3

Matches //; Conservative JF, mismatches 119, mismatches 187

QY 433 PLVSIYIPAYCANYIQRCVDSALNQTVVDLEVCICNDGSTDNTLEVINKLYGNPRVI 492

||||| : |||| | : : : |||| : : : | : |||| | : : ||| : ||| :

Db 241 QKTARNDL--QSMGFK-----TRFDSLE----YRQTKAAAYELPEKDL 278

Query Match 4.1%; Score 210; DB 1; Length 348;

[illegible]

Db 63 LAQAQNSGLIPSLINIGLDELAKSGGGGGYIARTDADDIASPQWIEKIVGEMEKDRSIIA 122

QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPFBSREKLTAMI-----AHHFRMFTIRAWHLT 594

Db 123 MGAWLEVLSEEDGRLARHHKHKIWKPKTRHEDIAAFAFPFCNPIHNTMIMRRS--VI 180

QY 595 DG---FNENIENAVDVMFKLSEVGFKHLNKCICNVRVLHGDNSTIK-----KLGI 643

Db 181 DGLRYDTERDWAEDYQFWYDVSGLGRLAYYPEALVKYRLHANOVSKSHSVRQHEITAQGI 240

QY 644 OK--KNHFVVVNSLNROGINNYNDKFDLDESRYIFNKTAEO--EEMDM 692

Db 241 QKTARNDFL--QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 14

US-08-878-360-11 Application US/08878360

; Sequence 11, Patent No. 5945322

; GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF

TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/878,360

FILING DATE: 18-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/683,426

FILING DATE:

APPLICATION NUMBER: 08/312,387

FILING DATE: September 26, 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-095B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-878-360-11

Query Match 4.1%; Score 210; DB 2; Length 348;

Best Local Similarity 26.4%; Pred. No. 4.1e-10;

Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLVSIYIPAYNCANYIQRQVDSALNQTVDLEVCICNDGSTDNTLVINKLYGNRRVRI 492

Db 3 PLVSVLCAYNVEKYFAQSLAAVYVNTWRNLILIVDDGSTDGTGLAIKDFQRDSRIKI 62

QY 493 MSK-PNGGIASASNAAV-SFAK-----GYVIGOLDSDDYLEPDAVELCLKEFLKDTLAC 545

Db 63 LAQAQNSGLIPSLINIGLDELAKSGGGGGYIARTDADDIASPQWIEKIVGEMEKDRSIIA 122

QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPFBSREKLTAMI-----AHHFRMFTIRAWHLT 594

Db 123 MGAWLEVLSEEDGRLARHHKHKIWKPKTRHEDIAAFAFPFCNPIHNTMIMRRS--VI 180

QY 595 DG---FNENIENAVDVMFKLSEVGFKHLNKCICNVRVLHGDNSTIK-----KLGI 643

Db 181 DGLRYDTERDWAEDYQFWYDVSGLGRLAYYPEALVKYRLHANOVSKSHSVRQHEITAQGI 240

QY 644 OK--KNHFVVVNSLNROGINNYNDKFDLDESRYIFNKTAEO--EEMDM 692

Db 241 QKTARNDFL--QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 15

US-08-478-140B-3

; Sequence 3, Application US/08478140B

; Patent No. 6127153

; GENERAL INFORMATION:

APPLICANT: JOHNSON, KARL F.

APPLICANT: ROTH, STEPHEN

APPLICANT: BUCZALA, STEPHANIE L.

TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO

TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A

TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,140B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Laura A. Coruzzi

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7188-017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-478-140B-3

Query Match 4.1%; Score 210; DB 3; Length 348;

Best Local Similarity 26.4%; Pred. No. 4.1e-10;

Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLVSIYIPAYNCANYIQRQVDSALNQTVDLEVCICNDGSTDNTLVINKLYGNRRVRI 492

Db 3 PLVSVLCAYNVEKYFAQSLAAVYVNTWRNLILIVDDGSTDGTGLAIKDFQRDSRIKI 62

QY 493 MSK-PNGGIASASNAAV-SFAK-----GYVIGOLDSDDYLEPDAVELCLKEFLKDTLAC 545

Db 63 LAQAQNSGLIPSLINIGLDELAKSGGGGGYIARTDADDIASPQWIEKIVGEMEKDRSIIA 122

QY 546 V--YTTNRNVNPDGSLIA-----NGYNWPFBSREKLTAMI-----AHHFRMFTIRAWHLT 594

```

Db      123  MGAWLEVLSEKOGNRLRAHHKHGKIWKXPTRHEDIATAEPFGPNLHNNTMIMRRS--VI 180
Qy      595  DG---FNPENTENAVDYDMFLKLEVCKFKHLNKICVNRVLHGDTISK-----KLGI 643
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      181  DGLGRYTERMDAEQFYFDVDSKGLRAYYPALVKYRUHANOVSSKHSVRQHETAQGI 240
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy      644  QK--KNHFVVVVNSLNQQGINYYNDKFDDLDERSKYIFNKTAEQY--EEMDM 692
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      241  OKTARNDFL---OSMGFK-----TFQSLE----YRQTAAAYELPEKDL 278
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

Search completed: January 4, 2003, 02:29:20
Job time : 71 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:26:45 ; Search time 203 Seconds
(without alignments)
90.090 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKGENIPVKNFINSITL 965

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues

Total number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pdb.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pdb.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	486.5	88.2	972	9	US-09-879-959-10
2	276.5	5.4	706	10	Sequence 10, Appl
3	276.5	5.4	715	10	Sequence 4950, Ap
4	246	4.8	322	10	Sequence 10511, A
5	224	4.4	332	10	Sequence 34, Appl
6	223.5	4.4	278	10	Sequence 22, Appl
7	223.5	4.4	674	10	Sequence 36, Appl
8	217.5	4.3	332	10	Sequence 200, App
9	210	4.1	348	12	Sequence 21, Appl
10	210	4.1	348	12	Sequence 3, Appl
11	209	4.1	337	12	Sequence 11, Appl
12	209	4.1	337	12	Sequence 5, Appl
13	208.5	4.1	270	10	Sequence 12, Appl
14	207	4.1	301	10	Sequence 39, Appl
15	205	4.0	389	10	Sequence 27, Appl
16	200.5	3.9	303	10	Sequence 34, Appl
17	199	3.9	297	10	Sequence 29, Appl
18	194	3.8	120	10	Sequence 31, Appl
19	194	3.8	322	10	Sequence 52, Appl
					Sequence 35, Appl

20	189.5	3.7	120	10	US-09-767-041-51	Sequence 51, Appl
21	187	3.7	150	10	US-09-924-358-29	Sequence 29, Appl
22	177	3.5	274	9	US-09-738-626-3905	Sequence 3905, Ap
23	176.5	3.5	345	9	US-09-738-626-3894	Sequence 3894, Ap
24	173.5	3.4	281	10	US-09-765-272-196	Sequence 196, App
25	168.5	3.3	210	10	US-09-767-041-47	Sequence 47, Appl
26	167.5	3.3	270	10	US-09-765-272-198	Sequence 198, App
27	165	3.2	313	10	US-09-900-038A-1	Sequence 1, Appl
28	164	3.2	573	10	US-09-815-242-12474	Sequence 12474, A
29	155	3.0	418	10	US-09-816-028A-33	Sequence 33, Appl
30	154.5	3.0	251	9	US-09-738-626-4202	Sequence 4202, Ap
31	153	3.0	303	10	US-09-765-272-202	Sequence 202, App
32	150	2.9	358	10	US-09-815-242-5714	Sequence 5714, Ap
33	145.5	2.9	589	12	US-10-001-851-26	Sequence 26, Appl
34	142	2.8	578	12	US-10-074-527-8	Sequence 8, Appl
35	138.5	2.7	1427	10	US-09-801-368-354	Sequence 354, App
36	137.5	2.7	187	9	US-09-973-457-4	Sequence 4, Appl
37	137.5	2.7	187	10	US-09-815-028-7	Sequence 7, Appl
38	137.5	2.7	187	12	US-10-074-527-4	Sequence 4, Appl
39	137.5	2.7	579	12	US-10-001-851-29	Sequence 29, Appl
40	137	2.7	1040	10	US-09-864-761-38325	Sequence 38325, A
41	137	2.7	1041	9	US-09-978-295A-498	Sequence 498, App
42	137	2.7	1041	9	US-09-978-697-498	Sequence 498, App
43	137	2.7	1041	9	US-09-978-192A-498	Sequence 498, App
44	137	2.7	1041	9	US-09-999-832A-498	Sequence 498, App
45	133.5	2.6	872	9	US-09-843-676-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-879-959-10
; Sequence 10, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: Deangelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESS
; TITLE OF INVENTION: IN BACILLUS SUBTILIS
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; PRIORITY FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 972
; TYPE: PRT
; ORGANISM: pasteurella multocida
US-09-879-959-10

Query Match	88.2%	Score	4486.5	DB	9	Length	972
Best Local Similarity	86.6%	Pred. No.	0	Mismatches	60	Indels	7
Matches	842	Conservative	63	Gaps	2		
Qy	1	MNTLSQAIKAYNSNDYELALKLFKSAEYGRKIVFQIKCKEL- ---STNS- ---YVS	53				
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAEYGRKIVFQIKCKEL- ---STNS- ---YVS	60				
Qy	54	EDKKNVCSDDIATQLLSNVYKLTLSSEKNSLKNKWSITGKKSENAEIRKVELVP	113				
Db	61	KEEVNVCSDPLDIATQLLSNVYKLTLSSEKNSLKNKWSITGKKSENAEIRKVELVP	120				
Qy	114	KDFPKDLVLAFLPDHVNDFTWYKRNKKSGLGIPVKNIGLSIIPTFNSRILDTLACL	173				
Db	121	KDFPKDLVLAFLPDHVNDFTWYKRNKKSGLGIPVKNIGLSIIPTFNSRILDTLACL	180				
Qy	174	VNQKNTYPEVVVADGSKENLLTIVQKVEQKLDIKYVRQKDYGLCAVRNGLRTAKY	233				

```
Db 181 VNKQTHPEVITDQSGEDLSPIIQYENKLDIRYVRQKONGFOASAARNGLRKY 240
QY 234 DFVSIILDCDMPAQOLWVSHYLFELLEDNDIVLIGPKYVDTHNITAEQFLNDPLYTESLP 293
Db 241 DFTGLDCDMPANPLWVSHVSAELDEDDLTITIGPKYIDTQHIDPKDFELNANSLLESPL 300
QY 294 ETATNNPSTSGNITSLDWRLEHFKKTONLRCLDSPPFRYFVAGNVAFAKWLKNGWFD 353
Db 301 EVKTNNSVAAGGEGTSLDWRLEQFEKTENLRSDSPFFFAAGNVAFAKWLKNGSGFPD 360
QY 354 EEFNHWGGEDVEFGYRLFAKCGCFRRVDGMAIHQBPPOKENTEREAGKSTITLKIVKEK 413
Db 361 EEFNHWGGEDVEFGYRLFRYGSFEKTDGIMAYHQBPPOKENTEREAGKSTITLDMREK 420
QY 414 VPIYRKLLPIEDSHIRPLVSIYPAYNCANYIQRCDVSALNQTVVDEVICNDGST 473
Db 421 VPIYRKLLPIEDSHIRNRPVLSIYPAYNCANYIQRCDVSALNQTVVDEVICNDGST 480
QY 474 DNTLEVINKLYGNPRVRIMSKPNGGIASNAVAFAKGYIIGQDSDSDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGIASNAVAFAKGYIIGQDSDSDYLEPDAVELC 540
QY 534 LKPEFLDKDTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTATTAMIAHPRMFTIRAWHL 593
Db 541 LKPEFLDKDTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTATTAMIAHPRMFTIRAWHL 600
QY 594 TDGFENENIENAVDYMFLKLSVCGKPKHLNKCYNRVLHGDNTSIRKLGTOKKNHFFVVN 653
Db 601 TDGFENKIENAVDYMFLKLSVCGKPKHLNKCYNRVLHGDNTSIRKLGTOKKNHFFVVN 660
QY 654 QSLNRQGITNYNDKFDLDESRYIFNKTAQYQEEEMDLKDLKIQNDKAKIAVSIFYP 713
Db 661 QSLNRQGITNYNDKFDLDESRYIFNKTAQYQEEEDILDKIQNDKAKIAVSIFYP 720
QY 714 NTLNGLVKLNNIEYNNKTFVILHVDKNHLPDTPDKKEILAFYHKHQVNILLNNDISYY 773
Db 721 NTLNGLVKLNNIEYNNKTFVILHVDKNHLPDTPDKKEILAFYHKHQVNILLNNDISYY 780
QY 774 TSNRLIKTEAHLNINKLSOLNCEYIIFDNHDSLFVKNDVAYMKKYDVGWGFNSALTH 833
Db 781 TSNRLIKTEAHLNINKLSOLNCEYIIFDNHDSLFVKNDVAYMKKYDVGWGFNSALTH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNDRSMVNGKASQGMFKYALPHELITTIKEVITSCQS 893
Db 841 DWIEKINAHPPFKLIKTYFNDNDRSMVNGKASQGMFKYALPHELITTIKEVITSCQS 900
QY 894 IDSVPENTEDIMPOFALLILEKKTGHVFNKSTLTTPMPWERKLQWNEQIQSAGKGENI 953
Db 901 IDSVPENTEDIMPOFALLILEKKTGHVFNKSTLTTPMPWERKLQWNEQIQSAGKGENI 960
QY 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972
```

RESULT 2

US-09-815-242-4950

Sequence 4950, Application US/09815242

Patent No. US20020061569AL

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA 011a

CURRENT APPLICATION NUMBER: US/09/815,242

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4950
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-815-242-4950
```

Query Match

5.4%; Score 276.5; DB 10; Length 706;

Best Local Similarity 21.5%; Pred. No. 1.2e-12; Indels 167; Gaps 25;
Matches 151; Conservative 115; Mismatches 268;

```
QY 260 DNDIVLIGPKYVD--THNITAEQFLNDPYLIESLPETATNNPSTSGNISLDMR--- 314
Db 2 NEDIKIVFDSIYRDKATNNLTITGVALDTTKES--PTFTINNENQVSAYNQIRVLREDV 59
QY 315 -----LEHFKKTDNLRCLDSPPFRYFVAGNVAFAKWLK-----VGWF 352
Db 60 NQIYQTEPAIEAGVVTLEGIKQKKVL-----PEHFOSSAHVVTVDPLNKKYVIPGTE 114
QY 353 DEEFNHWGGEDVEFGYRLFAKCGCFRRVIDGMAIHQBPPOKENTEREAGKSTITLKIVKE 412
Db 115 DKVTRLW-----IKAKGFKYMAKNGIS--HTQRAKIEKLRNQA--SYLNWLARN 161
QY 413 KVPYRKLLPIEDSHIRPLVSIYPAYNC--ANYIQRCDVSALNQTVVDEVICNDG 471
Db 162 EVLDI--EAMTQETATFHYQPKISAMPVYNVEEKWLRCLDSILNQVYTNWELCMADDA 219
QY 472 STD--NTLEVINKLYGNPRVRIM--SKPNGGIASNAVAFAKGYIIGQDSDSDYLEPDA 529
Db 220 STDPNVKKILTEYQQLDERIRVVRPEQNGHISEATNSALAIATGEFVALLDNDDELAINA 279
QY 530 VELCLKEFLDKDTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTATTAMIAHPRMFTIR 589
Db 280 FYEVVKVLNENPELDLIYSDEDIKIDMGNRSDDPAFK--PDWSPDLLLGTNYTSHLGVFRRS 338
QY 590 AWHLTDCGFENENIENAVDYMFLKLS-----VGKFKHLNKCYNRVLHGDNTSIRK--LGIQ 644
Db 339 ILEIIGGFRKGYEGSQDYDLVLRTEKTRITHIPKLYVYWRMLPTSTAVDQGSKGYA 398
QY 645 KKNHFVVVNVQSLNRQGINYN-----YDKFDDDESRYIFNKTAQYQEEEMDLKDL 696
Db 399 FEAGLRVQDALVRRGINGHATHGAANGLYDVYDI--ESER----- 438
QY 697 KLIQNDKAKIAVSIFYPNTLNG---LVKLNLIENKNIETVILHVDKNHLPDTPDKKEI 753
Db 439 -----LVSIILP--TKMGYKDVQRCVSSIEKTYQNYEIMADNGSTDPKMHELY 487
QY 754 LAFYHKHQVNILLNN--DISYVTSNRLIKTEAHLNINKLSOLNCEYIIFDNHDSLFVK 812
Db 488 AKFEQQLPGRFFVESIDIPF-----NFTSINNRAAKAHAGEVLLFLNNDTEVIT 536
QY 813 ND-----SYA-----YMKK 821
Db 537 ENWLTLMVSAFAQBQRICVGAKKLYPNNTVQHAGVILGLGVAGHGHYGYPHGDLGYFGR 596
QY 822 YDVGWGFNSALTHDWI-----EKINAHPPFKLIKTYFNDNDL 858
```

Db 597 LAINVNSAVTAACLLMKKADFDAVGGFEAFVAFNDVDL 637

RESULT 3

US-09-815-242-10511

; Sequence 10511, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10511

; LENGTH: 715

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-815-242-10511

Query Match 5.4%; Score 276.5; DB 10; Length 715;
Best Local Similarity 21.5%; Pred. No. 1.2e-12;
Matches 151; Conservative 115; Mismatches 268; Indels 167; Gaps 25;

Qy 260 DNDIVLIGPRKYD--THNITAEQPLNDPYLIESLPETATNNPSITSGNLSLWDR--- 314

Db 4 NEDIKVFIDSIYDRKATNNLTITGWALDTIKES--PTFTINNENQVSAYNIQRLVEDV 61

Qy 315 -----LEHFKKTDNRLCDSPRYFVAGNVAFSKEWLK-----VGWF 352

Db 62 NOIYOTEPALIEAGFVVTLEGIKOKVL-----PFFQSASHVTVDFPLNKYPVPGTE 116

Qy 353 DEEFNHWGGEDVEFGYRLFAGKCFPRVIGDMAIHQEPGPKENETEREAGKSITLKIVKE 412

Db 117 DKVTRLW-----IKAKGFKYMAKNGIS-HTIQRAKIEKLRNAQ--SYLNLWLRN 163

Qy 413 KVPYIYRKLPLPDSHIIRIPLVSIYIPAYNC-ANYIQRVDVSALNQTVDVLEVCINDG 471

Db 164 EYLDI--EAMTQEIATFIYQKISTAMPVYINVEEKWLRLCIDSIILNQVYNWELCMADDA 221

Qy 472 STD-NTLEVINKLYGNPNVRIM--SKPNGGIASANAASVPAKGYIIGOLDSDDVLEPDA 529

Db 222 STDPNVKILTEYQQLDERIRVVFREQHGHISEATNSALAIATGEFVALDNDDELAINA 281

Qy 530 VELCLKEFLKDKTLACVYTTNNRNPDSLIANGNWPFSREKLTMTAMIAHHFRMTFR 589

Db 282 FVEYVKVLNENPELDLIYSDEKIDMDGNRSDPAFK-PDWSDDLGLGTNYISHLGVYRS 340

Qy 590 AWHLTDGNGENENAVDYMFLKLSF-----VGKFKHLNKICYNRVLHGDNSTSIKK--LGIQ 644

Db 341 ILEEIGGFRKGYEGSQDYDLVLRFTTEKTKRITHIPKLVLYWRMLPTSTAVDQSGKGYA 400

Qy 645 KKNHFVVVQNSLNROGINYN-----YDKFDDLDSESKYIFNKTAETAEQOEMDKDL 696

Db 401 FEAGLRVAVODALVRGINGHATHGAANGLYDVIYDI-ESEK----- 440

Qy 697 KLIQNKDAKIAVSIFYPNTLNG---LVKKLNIIIEYNKNIFVIIHVDKNHLTPDIKKEI 753

Db 441 -----LVSIIIP-TKNGYKDVQRCVSSIIETTYQNYVEIINADNGSTDPKMHLY 489

Qy 754 LAFYHKHQVNILLNN-DISYYSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVK 812

Db 490 AKFEQQLPGRFEVESIDIPF-----NFSTINNRAAKKAHGEYLLFLNNDTEVIT 538

Qy 813 ND-----SYA-----YMKK 821

Db 539 ENWLTLMSVFAOQERIGCVGAKLLYPNNVQHVAGVILGGVAGHGHYGHGDLGYFGR 598

Qy 822 YDVGMNFSALTHDWI---EKINAHPPFPFKLIKTYFENDNL 858

Db 599 LAINVNSAVTAACLLMKKADFDAVGGFEAFVAFNDVDL 639

RESULT 4

US-09-767-041-34

; Sequence 34, Application US/09767041

; Patent No. US20020055168A1

; GENERAL INFORMATION:

; APPLICANT: Smith, Hilda

; TITLE OF INVENTION: STREPTOCOCCUS SUIIS VACCINES AND DIAGNOSTIC TESTS

; FILE REFERENCE: 2183-4726

; CURRENT APPLICATION NUMBER: US/09/767,041

; CURRENT FILING DATE: 2001-01-22

; PRIOR APPLICATION NUMBER: PCT/NL99/00460

; PRIOR FILING DATE: 1999-07-19

; PRIOR APPLICATION NUMBER: EP98202465.5

; PRIOR FILING DATE: 1998-07-22

; PRIOR APPLICATION NUMBER: EP98202467.1

; PRIOR FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 34

; LENGTH: 322

; TYPE: PRT

; ORGANISM: Streptococcus suis

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: CPSII

US-09-767-041-34

Query Match 4.8%; Score 246; DB 10; Length 322;
Best Local Similarity 25.3%; Pred. No. 7e-11;
Matches 92; Conservative 58; Mismatches 120; Indels 94; Gaps 15;

Qy 434 LVSIVIPAYNCANYIQRVDVSALNQTVDVLEVCINDGSDTNTLVINKLYGNPNVRIM 493

Db 4 LISVIVPIYVQDYLDKCIINSIINQYTNLEVLVNDGSDTDDSEKICLNKMGDRIKYY 63

Qy 494 SKPNGGIASANAASVPAKGYIIGOLDSDDVLEPDAVELCLKEFLKDKTLACVYTTNRV 553

Db 64 KKLINGLADARNFGLPEHATGKIAFVDSDDY-----IEVAMFERMHD-----NI 107

Qy 554 NPDGSLIA-----NGY-----NWPFSREKLTMTAMIAHHFRMTIRAHLLTDGF 597

Db 108 TEYNADIAIDFCLVDENGYYTKKRSNFHVLTRFETVKEFLS----- 150

Qy 598 NENIENAVDYMFLKLSYEVGKFKHLNKICYNRVLHGD-----NTSIKLGIOKNH-- 648

Db 151 GSNIEI-----NVWKLSYRDIIDKIKFQNNRSIGEDLLFNLEVLNNVTRVVVDTREYY 206

Qy 649 -FVVVQNSLNROGINYNYDKFDDLDSESKYIFNKTAETAEQOEMDKDLKLNQKAKIA 707

Db 207 NYVIRNSSLIQKFSINNIDLVRLE---NYPFKLREFSHYP---DAKVIREK----- 254

```
QY 708 VSIFYPNTLNGLVKKLNLIEN--KNIFVIIHVDKNHLT--PDIK-KEILAFYKHQV 762
DB 255 -----VKCLNKMYSTDCLDNEFLPILESYRKEIRRYFFIKAKRYLS--RKHLV 300
QY 763 NILL 766
DB 301 TLVL 304

RESULT 5
US-09-767-041-22
; Sequence 22, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUIIS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767, 041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Streptococcus suis
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: CPS2K
US-09-767-041-22

Query Match 4.4%; Score 224; DB 10; Length 332;
Best Local Similarity 24.4%; Pred. No. 3.1e-09;
Matches 81; Conservative 65; Mismatches 118; Indels 68; Gaps 14;

QY 435 VSIYIPAYNCANYIQRVDSALNQTVDVLEVCICNDGSTDNTLEVINKLYGNPRVRIMS 494
DB 4 ISIIPIYVNOYLSKCLNSIVNQTYSKHIEILLVNDGSTDNSEICLAYAKKDSRIRYFK 63
QY 495 KPNGGSIASNAVSFAKGYIGOLDSDYLEPAVELCLKEFLKDKTLACVYTNR--- 551
DB 64 KENGGLSDARNYGISRAKGDYLAFLIDSDDFIHSEFIQRLHEAIBRENALVAVAGYDRVDA 123
QY 552 -----NVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHLTDGFEENIEN 604
DB 124 SGHFLTAEPPLTNQAVLSGRN---VCKKLEA-DGHRF-----VVAWN-----KL 164
QY 605 VDYDMFLKSEVGKPKHLNKICYKRVLHGDNTSKKGIQKHNHFVVVYNQSLNRQGINYY 664
DB 165 YKKELDFRFEKGI-HEDEYFTYRLLY-----ELEKVAIVKE-----CLY 205
QY 665 NYDKFDLDES--RKYIFNKTAQYQ-BEMDML-----KDLKLIQNKDAKIAVSIF---Y 712
DB 206 YVDRENSIITSMTHRFHCLLEFQNERMDYFESRGDEL-LLECYSRFLAFVFLGKY 264
QY 713 PNTLNGLVKKLNLI-----IEYNKNIFVII 737
DB 265 NHWLSKQKQLQTLFRIVYKQLKQNKRLALLM 296

RESULT 6
US-09-767-041-36
; Sequence 36, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUIIS VACCINES AND DIAGNOSTIC TESTS
```

```
FILE REFERENCE: 2183-4726
CURRENT APPLICATION NUMBER: US/09/767,041
CURRENT FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: PCT/NL99/00460
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: EP98202465.5
PRIOR FILING DATE: 1998-07-22
PRIOR APPLICATION NUMBER: EP98202467.1
PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Patentin version 3.0
SEQ ID NO 36
LENGTH: 278
TYPE: PRT
ORGANISM: Streptococcus suis
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: CPS1K
US-09-767-041-36

Query Match 4.4%; Score 223.5; DB 10; Length 278;
Best Local Similarity 34.1%; Pred. No. 2.7e-09;
Matches 44; Conservative 29; Mismatches 55; Indels 1; Gaps 1;

QY 432 IPLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCICNDGSTDNTLEVINKLYGNPRVR 491
DB 4 ISKISIIPIYVNOYLSKCLNSIVNQTYSKHIEILLVNDGSTDNSEICLAYAKKDSRIR 63
QY 492 IMSKPNGGSIASNAVSFAKGYIGOLDSDYLEPAVELCLKEFLKDKTLACVYTNR 551
DB 64 YFKKENGGLSDARNYGISRAKGDYLAFLIDSDDFIHSEFIQRLHEAIBRENALVAVAGYDR 123
QY 552 NVNPDGSLI 560
DB 124 -VDASGHFL 131

RESULT 7
US-09-765-272-200
; Sequence 200, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Chol et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 674 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 200:
US-09-765-272-200

Query Match 4.4%; Score 223.5; DB 10; Length 674;
Best Local Similarity 20.1%; Pred. No. 9.1e-09;
Matches 96; Conservative 89; Mismatches 177; Indels 115; Gaps 17;

Qy 451 CVDASALNQTVDVLEVCICNDSTNTLEVKLYGNPRVRIMSPNGGIASASNAAYSF 510
Db 1 CLDSIITQYKNIEIVVNDGSDASGEICKEFSEMDHRILYIEQENAGLSAARNTGLNN 60
Qy 511 AKGYIGQDSDDYLEPAVELCLK---EFLKDKTLACVYTTNRN-----552
Db 61 MSGNTVTFVDSDDWTEQDYVETLYKKIYEQADIAVGNYYSFSESGMFYFHLGDSYFE 120
Qy 553 -VNPDGSLIANGYNWPEFSREKLTAMIAHHRMFTIRAWHL-----TDGFNENIE 602
Db 121 KYDYNVSIFENLYE---TQEMKSFALISAWGKLYKARLFEQLRFDIGKLGEDGY-----171
Qy 603 NAVDYDMFLKSEVGKFKHLNKCICY-NRVLHGD-----NTSIRKL 641
Db 172 --LNQKVYL-LSE--KVIYLNKSLVAYRIRKGLSRVWTEKWMHALVDMASERITLLANM 226
Qy 642 GIQKNHFVVNQSLNRQGINNYNDKFDLDESRYIFENKTAEOEMDMKLKLTQN 701
Db 227 GYPLEKHLAVRQMLEVSLAN---GQASGLSDTATY-----KEFEMKORLLNQLSRQBE 277
Qy 702 KDIAVSTFVPNTLGVKLNINIEYNKNIFFVILHVD-KNHLTPDICKKEILAF----756
Db 278 SEKKAIVLAANYGYVDVLTIKSCYHNRSIRFLIHSDFNEWIKQLNKLKLEFDSFI 337
Qy 757 ----YHKHQVNILLNNDISYTSNRLIKTEAHLNSKLNKLNCYIIFDNHDSLFVK 812
Db 338 INCRVTSEQIS-CYKSDISYTVFLRYFTADF---VOEDKALYLCDLVVTKNLDDLFAT 392
Qy 813 N-DSY-----AYMKYDVGCMNFSALTHDHIKIN 840
Db 393 DLQDYPLAARVDFGGRAYFGOEIFNAGVLLVNNAPFMKKENMTQKLDIVTNEWHDKVD 449

RESULT 8
US-09-767-041-21
; Sequence 21, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUITS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Streptococcus suis
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: CFS2J
US-09-767-041-21

Query Match 4.3%; Score 217.5; DB 10; Length 332;

Best Local Similarity 24.2%; Pred. No. 9.4e-09;
Matches 86; Conservative 67; Mismatches 116; Indels 87; Gaps 18;
Qy 435 VSIYIPAYNCANYIQRVCDSALNQTVDVLEVCICNDSTNTLEVKLYGNPRVRIMS 494
Db 4 VSIYIPFIETKYLRECLDSIIISQSYTNLEILLIDGSDSDTDICLEYAEOQDRIKFLR 63
Qy 495 .KPNGGIASASNAAVSFAGYIYGQDSDDYLEPDAVE---LCLKEFLKDKTLACVYTTNR 551
Db 64 LPNGGVSARNYCIKNTANYIMFVDSDDIVDGNIVESLYTCLKE--NDSLS-----114
Qy 552 NVNPDGSLIAN-GYNWPEFSREK-----LTTAMIAHHR--MFT-----IR 589
Db 115 -----GGLLATFDGNYQESELQCKIDLEIKERVDLGNENPNHYMSGIFNSPCKLYK 169
Qy 590 AWHLDGTFNENENAVDYDMFLKLS---EVGKFKHLNKCICY--NRVLHGDNTSIK-KLGI 643
Db 170 NIIYINGF--DTEQWLGEDLLENLYLNKIKKRVYVNRNLXFARRLSQSTTNTFKYDVFI 227
Qy 644 QKKN-----HFVVVQNQSLNRQGINNYNDKFDLDES-----KYI 679
Db 228 QLENLEKTFDLFVKIFGQYEFVFKETL-QWHIIYYSLLMEKNGDESLPKKLHIFKYL 286
Qy 680 FNKTAEOEMDML--KDLKLIONDKATAVSIFYPTNTLNLGVK-KLNIIIEYNKN 732
Db 287 YNR-----HSLDTLSIKRTSSVFKRICKLIVA-----NNLFKIFLNTLIREKN 330

RESULT 9
US-10-007-267-3
; Sequence 3, Application US/10007267
; Patent No. US20020127682A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0. Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,267
; FILING DATE: 03-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-007-267-11

Query Match 4.1%; Score 210; DB 12; Length 348;
Best Local Similarity 26.4%; Pred. No. 3.6e-08;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLSYIIPAYNCANYTORCVDSALNQTVDLEVCINCNDGSDTNTLEVINKLYGNPRVRI 492
DB 3 PLSVLICAYNVKYEYFAQSLAAVNTWRNLDLIYVDDGSDTGTDLAIAKDFQRDSRIKI 62
QY 493 MSK-PNGGIASNAAV-SFAK-----GYIGOLDSDDYLEPDAVELCLKEFLKDKTLAC 545
DB 63 LAQAQNSGLPLSLNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY 546 V--YTNRNVNPDGSLIA-----NGYNWPEFSREKLTAMI-----AHHFRMFTIRAWHLT 594
DB 123 MGAWLEVLSEKDGRLARHHKHKWKPKTRHEDIAAFAFPFGNPHNNTMIMRRS--VI 180
QY 595 DG---FNENTENAVDYMFLKLESEVGFKHLNKICYNRVHLHGNTSIK-----KLGI 643
DB 181 DGLRYDTERDWAEDYQFWYDVSKLGLAYYPEALVYRLHANOVSSKHSVRQHEIAQGI 240
QY 644 QK--KNHFVVVNSLNRQGINYNDFDLDLDESRYKIFNKTAIYQ--EEMDM 692
DB 241 QKTARNDL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 10

US-10-007-267-11
; Sequence 11, Application US/10007267
; Patent No. US20020127682A1

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,267

FILING DATE: 03-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/333,412

FILING DATE: 15-Jun-1999

APPLICATION NUMBER: 08/312,387

FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-10-007-267-11

Query Match

4.1%; Score 210; DB 12; Length 348;
Best Local Similarity 26.4%; Pred. No. 3.6e-08;
Matches 77; Conservative 57; Mismatches 110; Indels 48; Gaps 14;

QY 433 PLSYIIPAYNCANYTORCVDSALNQTVDLEVCINCNDGSDTNTLEVINKLYGNPRVRI 492
DB 3 PLSVLICAYNVKYEYFAQSLAAVNTWRNLDLIYVDDGSDTGTDLAIAKDFQRDSRIKI 62
QY 493 MSK-PNGGIASNAAV-SFAK-----GYIGOLDSDDYLEPDAVELCLKEFLKDKTLAC 545
DB 63 LAQAQNSGLPLSLNIGLDELAKSGGGGEYIARTDADDIASPGWIEKIVGEMEKDRSIIA 122
QY 546 V--YTNRNVNPDGSLIA-----NGYNWPEFSREKLTAMI-----AHHFRMFTIRAWHLT 594
DB 123 MGAWLEVLSEKDGRLARHHKHKWKPKTRHEDIAAFAFPFGNPHNNTMIMRRS--VI 180
QY 595 DG---FNENTENAVDYMFLKLESEVGFKHLNKICYNRVHLHGNTSIK-----KLGI 643
DB 181 DGLRYDTERDWAEDYQFWYDVSKLGLAYYPEALVYRLHANOVSSKHSVRQHEIAQGI 240
QY 644 QK--KNHFVVVNSLNRQGINYNDFDLDLDESRYKIFNKTAIYQ--EEMDM 692
DB 241 QKTARNDL---QSMGFK-----TRFDSLE-----YRQTKAAAYELPEKDL 278

RESULT 11

US-10-007-267-5

; Sequence 5, Application US/10007267

; Patent No. US20020127682A1

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,267

FILING DATE: 03-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/333,412

FILING DATE: 15-Jun-1999

APPLICATION NUMBER: 08/312,387

FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-10-007-267-5

Db 117 -----DAFVHRVTKQYRFKQDEVFNQKELEFLSKORH-----FCWSVWAKCFKKDI 165
QY 610 FLKLEVGKF-KHLN-----KICYNRVLHGDNTSIKKLGIOQKNHFFVNVQNSLRQGINV 663
Db 166 ILKSEFKIKIDERLNYGDEVLFCYIYFMFCEKIAVFTKCI----- 205
QY 664 YNDKEDDLDESRYIFNKTAEQEEMDKLKLQNKDAKTAVSIFY-PNTLNGLVKK 722
Db 206 YHYE-----FNPNGRY-----ENKNKEILNQNYHDKKKSNEIIRK 240
QY 723 LNNII---EYNKNIFVILHVD---KNHL 745
Db 241 LSKEFAHDEHOKLFEVLKKEEAGVKNRL 269
RESULT 14
US-09-816-028A-27
; Sequence 27, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: beta-1,3-galactosyltransferase from C. jejuni strain
; OTHER INFORMATION: OH4384 (ORF 6a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-09-816-028A-27
Query Match 4.1%; Score 207; DB 10; Length 301;
Best Local Similarity 21.0%; Pred. No. 4.9e-08;
Matches 78; Conservative 64; Mismatches 116; Indels 114; Gaps 13;
QY 435 VSIYIPAYNCANYIQRVCDSALNTQVVDLEVCICNDGSTDNTLEVINKLYGNPRVRIM- 493
Db 4 ISIILPTNVQXIARATESCINQTFKDIETIIVDDCGNDNSINIAKEYSKDKRIKIIH 63
QY 494 SKPNGGIASANAASFAGKYYIGQLSDDDYLDPAVELCLKEFLKDKTLACVYTTNRV 553
Db 64 NEKNLGLRLARYEGYKVANSPIYFLDPDDYLELNAECIKILDEQDEVDLVF----- 117
QY 554 NPDGSLIANGYNWPEFSREKLTMTAMIAHFRMTIRAWHLTDGFPNTE-NAVYDMELK 612
Db 118 -----KKFDNFSNGYSYKKEPVKKI-TAKKNLY-----FNAIVESNVIY----- 129
QY 613 LSEVGKFKHLNKICYNRVLHGDNTSIKKLGIOQKNHFFVNVQNSLRQGINVYNDKFDL 672
Db 130 -----KKFDNFSNGYSYKKEPVKKI-TAKKNLY-----WTMNGKL 162
QY 673 DESRYIFNKTAEQEEMDKLKLQNKDAKIAVSIFYPNTLNGLVKLNIIYEYKN 732
Db 163 IRKKLYL-----EAPASFLERLEKDVKNMAED-----VLIYIP-----MLSQAKIAYMCN 208
QY 733 IFVILHVDKHNLTDPDKKEILAFYKHQVNNILLNNDISYITSNRLIKTEAHLNSINK-L 791
Db 209 LY---HYVPNNNSICNTKNEVLVKNIIQELQLVLN---YLRQNYI-----LNKYC 252

QY 792 SQLNLNCEYIIF 803
Db 253 SVLYVLIKYLLY 264
RESULT 15
US-09-816-028A-34
; Sequence 34, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 34
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: glycosyltransferase of C. jejuni OH4384 (ORF 4a of
; OTHER INFORMATION: lipooligosaccharide (LOS) biosynthesis locus)
US-09-816-028A-34
Query Match 4.0%; Score 205; DB 10; Length 389;
Best Local Similarity 21.2%; Pred. No. 9.9e-08;
Matches 114; Conservative 81; Mismatches 146; Indels 196; Gaps 25;
QY 435 VSIYIPAYNCANYIQRVCDSALNTQVVDLEVCICNDGSTD-NTLEVINKLYGNPRVRIM 493
Db 4 IGVYIPIYVNBKYLECDSVINQTYNLEIILVNDGSTDDEHSLNIAKEYTLKDKRIITLF 63
QY 494 SKPNGGIASANAASFAGKYYIGQLSDDDYLDPAVELCLKEFLKDKTLACVYTTNRV 553
Db 64 DKNKGGLSSARNIGIEYFSGEY-----KLKNTOKHENSLEIF 102
QY 554 NPDGSLIANGYNWPEFSREKLTMTAMIAHFRMTIRAWHLTDGFPN-ENIEN----AVDY 607
Db 103 QLDGN---NPYN-----IYKAYKSSQAFNNEKDLTNFTYPSIDY 138
QY 608 DMFLKLEVGKFKHLNKICYNRVLHGDNTSIKKLGIOQKNHFFVNVQNSLRQGINVYND 667
Db 139 IIFLSDSNYKWL-----NCIEECVIRKMNVDVL-----WFDHD 171
QY 668 -KFDDLDESRYIFNKTAEQEEM-DMLKDLKLQNKDAKIAVSIFYPNTLNGLVKLN 725
Db 172 CTYED-----NIKKNHKKTRMEIFDFKECIITPEYANRALSVGSRDISFG----WNG 221
QY 726 IIEYNKNIFVILHVDKHNLTDPDKKEILAFYKHQVNNILLNNDIS-----YITSNRLIK 780
Db 222 MIDFN-----FLQIKLKF-----INFIINEDIHFGILFASANKIYV 259
QY 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDYVAYMKYDVGVMNFSALTDHTEKIN 840
Db 260 -----LSQKLYLCRLRANS-----ISNHD-----ITKAN 285
QY 841 AHPPPEKLIKTIYFNDNDRKSMN-VKASQSGMPMKYALPHELITIIKEVITTSQSIDSVPE 899
Db 286 VSEYFKDIYET-FGENAKEAKNYLKAASR-----VITALKLEIEFFKD 336
QY 900 YNTEDWFOFALLILEKTKGHVFNFTSTUTYMPWEKRLQWTFNEQIOSAKKGENIPVN 956
Db 327 QKNEN-----ALAIKE-----TFLPCYAK-----KALMIKKFKKDKPLN 359

Search completed: January 4, 2003, 02:49:25
Job time : 206 secs

THIS PAGE BLANK (USPTO)


```
QY 61 CDSLDIATQLLSNVKKITLSESEKNSLKNKWKSTGKSENAEIRKVELVPKDPKDL 120
Db 61 CDSLDIATQLLSNVKKITLSESEKNSLKNKWKSTGKSENAEIRKVELVPKDPKDL 120
QY 121 VLAPLDPHNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180
Db 121 VLAPLDPHNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180
QY 181 PFEVWVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
Db 181 PFEVWVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLVHSHYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPVLIESLPETATNN 300
Db 241 CDMAPOOLVHSHYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPVLIESLPETATNN 300
QY 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
Db 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360
QY 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHQEPGKENEETEREAGKSITLKIYKEKVPYIYRK 420
Db 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHQEPGKENEETEREAGKSITLKIYKEKVPYIYRK 420
QY 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTDNTLEVI 480
Db 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGIASANAASVFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGIASANAASVFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKLESEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHVVVNGSLNRQ 660
Db 601 IENAVDYDMFLKLESEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHVVVNGSLNRQ 660
QY 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDAKIAVSIYPNTLGLV 720
Db 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDAKIAVSIYPNTLGLV 720
QY 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780
Db 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780
QY 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
```

RESULT 2
PCT-US99-26501-3
; Sequence 3, Application PC/TUS9926501
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
; FILE REFERENCE: 617481-5

```
; CURRENT APPLICATION NUMBER: PCT/US99/26501  
; EARLIER FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,929  
; EARLIER FILING DATE: 1998-11-11  
; EARLIER APPLICATION NUMBER: 09/283,402  
; EARLIER FILING DATE: 1999-04-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 965  
; TYPE: PRT  
; ORGANISM: Pasteurella multocida  
PCT-US99-26501-3
```

```
Query Match 100.0%; Score 5089; DB 1; Length 965;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIVEFQIICKKELSTNSVSEDKNSV 60  
Db 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIVEFQIICKKELSTNSVSEDKNSV 60  
QY 61 CDSLDIATQLLSNVKKITLSESEKNSLKNKWKSTGKSENAEIRKVELVPKDPKDL 120  
Db 61 CDSLDIATQLLSNVKKITLSESEKNSLKNKWKSTGKSENAEIRKVELVPKDPKDL 120  
QY 121 VLAPLDPHNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180  
Db 121 VLAPLDPHNDFTWYKRNKKSIGIKPVNKNIGLSIIIPFNRSRIIDITLACLVNOKTNY 180  
QY 181 PFEVWVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240  
Db 181 PFEVWVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240  
QY 241 CDMAPOOLVHSHYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPVLIESLPETATNN 300  
Db 241 CDMAPOOLVHSHYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPVLIESLPETATNN 300  
QY 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360  
Db 301 PSITSKGNISLDRLEHFKKTDNLRCDSPFRYFVAGNVAFSKEWLNKVGWDFEENHWG 360  
QY 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHQEPGKENEETEREAGKSITLKIYKEKVPYIYRK 420  
Db 361 GEDVEFGYRLFAKGCFFRVIDGGMAIHQEPGKENEETEREAGKSITLKIYKEKVPYIYRK 420  
QY 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTDNTLEVI 480  
Db 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTDNTLEVI 480  
QY 481 NKLYGNPRVRIMSKPNGGIASANAASVFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540  
Db 481 NKLYGNPRVRIMSKPNGGIASANAASVFAKGYIIGQLSDDDYLEPDAVELCLKEFLKD 540  
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600  
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFEN 600  
QY 601 IENAVDYDMFLKLESEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHVVVNGSLNRQ 660  
Db 601 IENAVDYDMFLKLESEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHVVVNGSLNRQ 660  
QY 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDAKIAVSIYPNTLGLV 720  
Db 661 INYNYDKFDLDESRRKIFNKTAEQEEMDMKDLKLIQNKDAKIAVSIYPNTLGLV 720  
QY 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780  
Db 721 KKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISYTSNRLIK 780  
QY 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840  
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMNFSALTHDWEIKIN 840
```

Qy 841 AHPPEKKLIKTYFNDNLSMNVKASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
|||||
Db 841 AHPPEKKLIKTYFNDNLSMNVKASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
|||||
Qy 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKQWNTNEQIOSAKKGENIPVNFII 960
|||||
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKQWNTNEQIOSAKKGENIPVNFII 960
|||||
Qy 961 NSITL 965
|||||
Db 961 NSITL 965
|||||

RESULT 3
PCT-US01-13395-4
; Sequence 4, Application PC/TUS0113395
; GENERAL INFORMATION:
; APPLICANT: DE ANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND
; FILE REFERENCE: 618755-9/JP/199,538
; CURRENT APPLICATION NUMBER: PCT/US01/13395
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
PCT-US01-13395-4

Query Match 99.6%; Score 5067; DB 1; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDKKNSV 60
|||||
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDKKNSV 60
|||||

Qy 61 CDSSLDIATQLLSNVKLLTISESEKNSLKNKWSITCKSENAETRKVELVPKDFPKDL 120
|||||
Db 61 CDSSLDIATQLLSNVKLLTISESEKNSLKNKWSITCKSENAETRKVELVPKDFPKDL 120
|||||

Qy 121 VLAPLDHVNDFTWYKNRKKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACLNVQKTY 180
|||||
Db 121 VLAPLDHVNDFTWYKNRKKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACLNVQKTY 180
|||||

Qy 181 PFEVVVADGSKENLLTIVQKYEOKLDIKYVRQKDYQOLCAVRNLGLRTAKYDFVSILD 240
|||||
Db 181 PFEVVVADGSKENLLTIVQKYEOKLDIKYVRQKDYQOLCAVRNLGLRTAKYDFVSILD 240
|||||

Qy 241 CDMAPOQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPILIESPETATNN 300
|||||
Db 241 CDMAPOQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEQFLNDPILIESPETATNN 300
|||||

Qy 301 PSITSGNISLDWRLHEHFKKTDNLCDSPFRYFVAGNVAFSEKWLNVKGVDFEENHWG 360
|||||
Db 301 PSITSGNISLDWRLHEHFKKTDNLCDSPFRYFVAGNVAFSEKWLNVKGVDFEENHWG 360
|||||

Qy 361 GEDVEFGYRLAKGCFRVIDGMAIHQEPGKNETEREAGKSTLTKIVKEKVPYIYRK 420
|||||
Db 361 GEDVEFGYRLAKGCFRVIDGMAIHQEPGKNETEREAGKSTLTKIVKEKVPYIYRK 420
|||||

Qy 421 LLPEDSHHRIPLYSIIPAYNCANYIQRVDSALNQTVVDLEVICINDGSTDNTLEVI 480
|||||
Db 421 LLPEDSHHRIPLYSIIPAYNCANYIQRVDSALNQTVVDLEVICINDGSTDNTLEVI 480
|||||

Qy 481 NKLYGNNPRVIRMSKPNGGIASASNAVSFAKYYIGQLSDDDYLEPDAVELCLKEFLKD 540
|||||

Db 481 NKLYGNNPRVIRMSKPNGGIASASNAVSFAKYYIGQLSDDDYLEPDAVELCLKEFLKD 540
|||||

Qy 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHHPMTTIRAWHLTDGFEN 600
|||||

Db 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHHPMTTIRAWHLTDGFEN 600
|||||

Qy 601 IENAVDYDMFLKLSVKGFKHLNKCYNVNLHGDNITSIKKLGIOKKNHFFVNVNOSLNROG 660
|||||

Db 601 IENAVDYDMFLKLSVKGFKHLNKCYNVNLHGDNITSIKKLGIOKKNHFFVNVNOSLNROG 660
|||||

Qy 661 INYNYDKFDDLDDESRYIFNKTAEYOEEMDLKDLKIQNKDAKIAVSIFYPTNLGLV 720
|||||

Db 661 INYNYDKFDDLDDESRYIFNKTAEYOEEMDLKDLKIQNKDAKIAVSIFYPTNLGLV 720
|||||

Qy 721 KKLNNIIEYNNKIFVIIHVDKNHLTPDIKKEILAFYHKKHVNILLNNDISYTSNRLIK 780
|||||

Db 721 KKLNNIIEYNNKIFVIIHVDKNHLTPDIKKEILAFYHKKHVNILLNNDISYTSNRLIK 780
|||||

Qy 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAALTHDWEKIN 840
|||||

Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAALTHDWEKIN 840
|||||

Qy 841 AHPPEKKLIKTYFNDNLSMNVKASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
|||||

Db 841 AHPPEKKLIKTYFNDNLSMNVKASQGMFKYALPHELLTIIKEVITSCQIDSVPY 900
|||||

Qy 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKQWNTNEQIOSAKKGENIPVNFII 960
|||||

Db 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKQWNTNEQIOSAKKGENIPVNFII 960
|||||

Qy 961 NSITL 965
|||||

Db 961 NSITL 965
|||||

RESULT 4
PCT-US99-07289-3
; Sequence 3, Application PC/TUS9907289
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; FILE REFERENCE: 617022-7
; CURRENT APPLICATION NUMBER: PCT/US99/07289
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: 60/080,414
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 60/178,851
; EARLIER FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
PCT-US99-07289-3

Query Match 98.6%; Score 5017.5; DB 1; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLST-----NSYVS 53
|||||

Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDSYVS 60
|||||

Qy 54 EDKNSVCDSSLDIATQLLSNVKLLTISESEKNSLKNKWSITCKSENAETRKVELV 113
|||||

Db 61 EDKNSVCDSSLDIATQLLSNVKLLTISESEKNSLKNKWSITCKSENAETRKVELV 120
|||||

Qy 114 KDFPKDLVLAPLDHVNDFTWYKNRKKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACL 173
|||||

Db 121 KDFPKDLVLAPLDHVNDFTWYKNRKKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACL 180
|||||

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida (PmCS)
PCT-US99-07289-8

Query Match 98.6%; Score 5017.5; DB 1; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;

QY 1 MNTLSQAIAKAYNSNDYELALKLFPEKSAETVGRKIVVEQIIKCKEKLST-----NSVVS 53
Db 1 MNTLSQAIAKAYNCNDYELALKLFPEKSAETVGRKIVVEQIIKCKEKLSTNSVSDNSVS 60

QY 54 EDKNSVCDSSLDATATOLLISNVKKLTLSEKNSLKNKKWSITGKKSSENAEIRKVELVP 113
Db 61 EDKNSVCDSSLDATATOLLISNVKKLTLSEKNSLKNKKWSITGKKSSENAEIRKVELVP 120

QY 114 KDFPKDLVLAFLPDHVNDFTWYKNRXXSLGKIPVKNKNIIGLSIIITPTNRSRILDTIACL 173
Db 121 KDFPKDLVLAFLPDHVNDFTWYKNRXXSLGKIPVKNKNIIGLSIIITPTNRSRILDTIACL 180

QY 174 VNOKTNTPFVWVADDSKSNLLTIIVQYEQKLDIKYVROKDYGYOLCAVRNLGLRTAKY 233
Db 181 VNOKTNTPFVWVADDSKSNLLTIIVQYEQKLDIKYVROKDYGYOLCAVRNLGLRTAKY 240

QY 234 DFVSLDCCMAPQOLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 293
Db 241 DFVSLDCCMAPQOLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 300

QY 294 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 360

QY 354 EEFNHGGEDEVFGYRLFAKCGCFRVIDGMAIHQEPGKENETREAGSKITLKIYREK 413
Db 361 EEFNHGGEDEVFGYRLFAKCGCFRVIDGMAIHQEPGKENETREAGSKITLKIYREK 420

QY 414 VPYIRKLLPIEDSHIHRIPLVSIYIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 473
Db 421 VPYIRKLLPIEDSHIHRIPLVSIYIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 480

QY 474 DNTLEVINKLKGNPRVRIMSKPNNGSIASASNAAYSAFAGYIGOLDSDDYLEPDAVELC 533
Db 481 DNTLEVINKLKGNPRVRIMSKPNNGSIASASNAAYSAFAGYIGOLDSDDYLEPDAVELC 540

QY 534 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600

QY 594 TDGFNENIENAVDYDMFLKSEVGKPKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 653
Db 601 TDGFNENIENAVDYDMFLKSEVGKPKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 660

QY 654 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 713
Db 661 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 720

QY 714 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISY 773
Db 721 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISY 780

QY 774 TSNRLIKTEAHLNINLKSNLQNLNCEYIIFDNHDSLFVKNDSYAYKKYDVGWNFSAETH 833
Db 781 TSNRLIKTEAHLNINLKSNLQNLNCEYIIFDNHDSLFVKNDSYAYKKYDVGWNFSAETH 840

QY 834 DWIEKINAHPPPKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQS 893
Db 841 DWIEKINAHPPPKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQS 900

QY 894 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLYMPWERKLOWTNEQIOSAKKGENI 953
Db 954 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLYMPWERKLOWTNEQIOSAKKGENI 960

QY 174 VNOKTNTPFVWVADDSKSNLLTIIVQYEQKLDIKYVROKDYGYOLCAVRNLGLRTAKY 233
Db 181 VNOKTNTPFVWVADDSKSNLLTIIVQYEQKLDIKYVROKDYGYOLCAVRNLGLRTAKY 240

QY 234 DFVSLDCCMAPQOLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 293
Db 241 DFVSLDCCMAPQOLWVHSYLTLELNDIVLIGPRKYVDTHNTAEQFLNDPYLIESLP 300

QY 294 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPSTISKGNISLDWRLEHFKTDNLRLCDSPPRYFVAGNVAFSKWLNKVGWFD 360

QY 354 EEFNHGGEDEVFGYRLFAKCGCFRVIDGMAIHQEPGKENETREAGSKITLKIYREK 413
Db 361 EEFNHGGEDEVFGYRLFAKCGCFRVIDGMAIHQEPGKENETREAGSKITLKIYREK 420

QY 414 VPYIRKLLPIEDSHIHRIPLVSIYIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 473
Db 421 VPYIRKLLPIEDSHIHRIPLVSIYIPAYNCANYTORCVDSALNOTVVDLEVCIKNDGST 480

QY 474 DNTLEVINKLKGNPRVRIMSKPNNGSIASASNAAYSAFAGYIGOLDSDDYLEPDAVELC 533
Db 481 DNTLEVINKLKGNPRVRIMSKPNNGSIASASNAAYSAFAGYIGOLDSDDYLEPDAVELC 540

QY 534 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTFLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600

QY 594 TDGFNENIENAVDYDMFLKSEVGKPKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 653
Db 601 TDGFNENIENAVDYDMFLKSEVGKPKHLNKICYNRVLHGDNTSIKKLGIOKKNHFVVYN 660

QY 654 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 713
Db 661 QSLNRQGINYNVDKFDLDESRYIFNKTAEOEEMDMKDLKLIQNKDKAKIAVSIFYP 720

QY 714 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISY 773
Db 721 NTLNGLVKKLNIIIEYNKNIIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISY 780

QY 774 TSNRLIKTEAHLNINLKSNLQNLNCEYIIFDNHDSLFVKNDSYAYKKYDVGWNFSAETH 833
Db 781 TSNRLIKTEAHLNINLKSNLQNLNCEYIIFDNHDSLFVKNDSYAYKKYDVGWNFSAETH 840

QY 834 DWIEKINAHPPPKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQS 893
Db 841 DWIEKINAHPPPKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSQS 900

QY 894 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLYMPWERKLOWTNEQIOSAKKGENI 953
Db 954 IDSVPEYNTEDIWFQFALLILEKKTGHVFNKTTSTLYMPWERKLOWTNEQIOSAKKGENI 960

RESULT 5
PCT-US99-07289-8
; Sequence 8, Application PC/TUS9907289
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 617022-7
; CURRENT APPLICATION NUMBER: PCT/US99/07289
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: 60/080,414
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 60/178,851
; EARLIER FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29

Db 901 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSTLTYPMPWERKLOWTNEQIOSAKKGENI 960
Qy 954 PVNKFIINSITL 965
Db 961 PVNKFIINSITL 972
RESULT 6
US-09-283-402-3
; Sequence 3, Application US/09283402
; GENERAL INFORMATION:
; APPLICANT: Paul DeAngelis
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA
; FILE REFERENCE: 5820.530
; CURRENT FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-283-402-3
Query Match 98.6%; Score 5017.5; DB 16; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLST-----NSVYS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSVYSEDNSVYS 60
Qy 54 EDKNSVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 113
Db 61 EDKNSVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 120
Qy 114 KDFPKDLVLA PLDPHVDFTWYKRRKSLGIPKVNKNIGLSIIPTFNRSLDITLACL 173
Db 121 KDFPKDLVLA PLDPHVDFTWYKRRKRLGIPKVNKNIGLSIIPTFNRSLDITLACL 180
Qy 174 VNQKTNPFVAVVADGSGKENLLTIQVKEOKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTNPFVAVVADGSGKENLLTIQVKEOKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSLDCDMAPOQLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSLDCDMAPOQLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPSITSGNISLDWRLEHFKKTDNLRLCDSPPRYFVAGNVAFSKENLKNVGFWD 353
Db 301 ETATNNPSITSGNISLDWRLEHFKKTDNLRLCDSPPRYFVAGNVAFSKENLKNVGFWD 360
Qy 354 EEFNHWGDEVEFYRLFAKCGCFPRIDGMAIHQEPGKKNETEREAGSKITLKIIVEK 413
Db 361 EEFNHWGDEVEFYRLFAKCGCFPRIDGMAIHQEPGKKNETEREAGSKITLKIIVEK 420
Qy 414 VPYIRKLLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALAQTVVDVLEVCINCDSGT 473
Db 421 VPYIRKLLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALAQTVVDVLEVCINCDSGT 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAVSFAGYIYGOLDSDDDYLEDPAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAVSFAGYIYGOLDSDDDYLEDPAVELC 540
Qy 534 LKEFLKOKTTLACVYTTNRNVPDGLIANGYNWPEFSREKLTUTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKOKTTLACVYTTNRNVPDGLIANGYNWPEFSREKLTUTAMIAHFRMFTIRAWHL 600
Qy 594 TDGNEINENAVDYDMFLKSEVCKFKHLNKICYNRVLHGDNSTIKKLGIOKKNHFVVVN 653
Db 601 TDGNEINENAVDYDMFLKSEVCKFKHLNKICYNRVLHGDNSTIKKLGIOKKNHFVVVN 660

Qy 654 OSLNRQGINYNYDKFDDDESRYKIFNKTAEYQEEMDKLDKLIQNKDAKIAVSIFYP 713
Db 661 OSLNRQGINYNYDKFDDDESRYKIFNKTAEYQEEDMLKDLKLIQNKDAKIAVSIFYP 720
Qy 714 NTLNGLVKKLNNIIIEYNKNIFVIIHVDKNHLLTPDIKKELAFYHKHQNVLNNDISY 773
Db 721 NTLNGLVKKLNNIIIEYNKNIFVIIHLDKNHLLTPDIKKELAFYHKHQNVLNNDISY 780
Qy 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWNFSA LTH 833
Db 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWNFSA LTH 840
Qy 834 DWIEKINAHPPPKLIKITYFNDNDRSMNVKGASOGMFMKYALPHELLFTIIEKITSCOS 893
Db 841 DWIEKINAHPPPKLIKITYFNDNDRSMNVKGASOGMFMKYALRHALLTIIEKITSCOS 900
Qy 894 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSTLTYPMPWERKLOWTNEQIOSAKKGENI 953
Db 901 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSTLTYPMPWERKLOWTNEQIOSAKKGENI 960
Qy 954 PVNKFIINSITL 965
Db 961 PVNKFIINSITL 972
RESULT 7
US-09-283-402-8
; Sequence 8, Application US/09283402
; GENERAL INFORMATION:
; APPLICANT: Paul DeAngelis
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA
; FILE REFERENCE: 5820.530
; CURRENT FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida (PmCS)
US-09-283-402-8
Query Match 98.6%; Score 5017.5; DB 16; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLST-----NSVYS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETYGRIKIVFEQIIKCKEKLSTNSVYSEDNSVYS 60
Qy 54 EDKNSVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 113
Db 61 EDKNSVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVP 120
Qy 114 KDFPKDLVLA PLDPHVDFTWYKRRKSLGIPKVNKNIGLSIIPTFNRSLDITLACL 173
Db 121 KDFPKDLVLA PLDPHVDFTWYKRRKRLGIPKVNKNIGLSIIPTFNRSLDITLACL 180
Qy 174 VNQKTNPFVAVVADGSGKENLLTIQVKEOKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTNPFVAVVADGSGKENLLTIQVKEOKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSLDCDMAPOQLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSLDCDMAPOQLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPSITSGNISLDWRLEHFKKTDNLRLCDSPPRYFVAGNVAFSKENLKNVGFWD 353
Db 301 ETATNNPSITSGNISLDWRLEHFKKTDNLRLCDSPPRYFVAGNVAFSKENLKNVGFWD 360
Qy 354 EEFNHWGDEVEFYRLFAKCGCFPRIDGMAIHQEPGKKNETEREAGSKITLKIIVEK 413
Db 354 EEFNHWGDEVEFYRLFAKCGCFPRIDGMAIHQEPGKKNETEREAGSKITLKIIVEK 413

Db	361	EEFNHGGEDVEFGYRLFPKGCFFRVIDGGMAYHQPPGKENETEREAGKSITLKIVKEK	420
Qy	414	VPYIYKLLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVVLEVCICNDGST	473
Db	421	VPYIYKLLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVVLEVCICNDGST	480
Qy	474	DNTLEVINLKYGNPRVRIMSKPNGGIASASNAVSAFAGYIIGQLSDSDYLEPDAVELC	533
Db	481	DNTIEVINLKYGNPRVRIMSKPNGGIASASNAVSAFAGYIIGQLSDSDYVLEPDAVELC	540
Qy	534	LKEFLDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL	593
Db	541	LKEFLDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL	600
Qy	594	TGDFENENIENAVDYMFLKSEVGFYHFKHNLKICYNRVLHGDNTSIKLGIOKKNHFVVVN	653
Db	601	TGDFENENIENAVDYMFLKSEVGFYHFKHNLKICYNRVLHGDNTSIKLGIOKKNHFVVVN	660
Qy	654	QSLNRQGINYNDKFDLDESRYIFNKTAEYQEEEDMLDKLQIONKDAKIAVSIFYP	713
Db	661	QSLNRQGINYNDKFDLDESRYIFNKTAEYQEEEDMLDKLQIONKDAKIAVSIFYP	720
Qy	714	NTNLGLVKLLNNIIEYNKNIFVILHLDKNHLPDIKKEILAFYHKKHQNILLNNDISYY	773
Db	721	NTNLGLVKLLNNIIEYNKNIFVILHLDKNHLPDIKKEILAFYHKKHQNILLNNDISYY	780
Qy	774	TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETH	833
Db	781	TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETH	840
Qy	834	DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQCS	893
Db	841	DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQCS	900
Qy	894	IDSVPYNTEDIWFQALLILEKKTGHVNKTSTLTTPWPERKLOWNEQIOSAKKGENI	953
Db	901	IDSVPYNTEDIWFQALLILEKKTGHVNKTSTLTTPWPERKLOWNEQIOSAKKGENI	960
Qy	954	PVNKFIINSITL 965	
Db	961	PVNKFIINSITL 972	
RESULT 8			
US-09-469-200-3			
; Sequence 3, Application US/09469200			
; GENERAL INFORMATION:			
; APPLICANT: Board of Regents of the University of Oklahoma			
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS			
; TITLE OF INVENTION: OF USE			
; FILE REFERENCE: 617022-7			
; CURRENT APPLICATION NUMBER: US/09/469,200			
; CURRENT FILING DATE: 1999-12-21			
; PRIOR APPLICATION NUMBER: 60/080,414			
; PRIOR FILING DATE: 1998-04-02			
; PRIOR APPLICATION NUMBER: 60/178,851			
; PRIOR FILING DATE: 1998-10-26			
; NUMBER OF SEQ ID NOS: 29			
; SOFTWARE: Patent in Ver. 2.0			
; SEQ ID NO 3			
; LENGTH: 972			
; TYPE: PRT			
; ORGANISM: Pasteurella multocida			
US-09-469-200-3			
Query Match 98.6%; Score 5017.5; DB 18; Length 972;			
Best Local Similarity 98.0%; Pred. No. 0;			
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;			
Qy	1	MNTLSQAIKAYNSNDYELAKLFEKSAETYGRIKVEFQIKCKEKLST-----NSYVS	53
Db	1	MNTLSQAIKAYNSNDYELAKLFEKSAETYGRIKVEFQIKCKEKLSTNSYVSEDNSYVS	60

Qy	54	EDKKNVCDSSLDIAQALLSNVKKLTSESEKSNLKNWKSITGKSKSENAEIRKVELVP	113
Db	61	EDKKNVCDSSLDIAQALLSNVKKLTSESEKSNLKNWKSITGKSKSENAEIRKVELVP	120
Qy	114	KDPKDLVLAPLDPHNDFTWYKRNKSLGKIPVKNIGLSIIPTFNRSRIIDITLACL	173
Db	121	KDPKDLVLAPLDPHNDFTWYKRNKSLGKIPVKNIGLSIIPTFNRSRIIDITLACL	180
Qy	174	VNOKTWPPEVVVADGSGKENLTIIVOKYEQKLDIKVVRQKDYQYQCALVNRNIGLRTAKY	233
Db	181	VNOKTWPPEVVVADGSGKENLTIIVOKYEQKLDIKVVRQKDYQYQCALVNRNIGLRTAKY	240
Qy	234	DFYSIILDCDMAPQOLWVHSYLTLELLEDNDIVLIGPRKYVDTHTNITABQFLNDPILIESLP	293
Db	241	DFYSIILDCDMAPQOLWVHSYLTLELLEDNDIVLIGPRKYVDTHTNITABQFLNDPILIESLP	300
Qy	294	ETATNNPSTSGNISLDMRLEHFKKTDNLRLCDSFPRYFVAGNVAFSKEWLNKVGWFD	353
Db	301	ETATNNPSTSGNISLDMRLEHFKKTDNLRLCDSFPRYFVAGNVAFSKEWLNKVGWFD	360
Qy	354	EEFNHGGEDVEFGYRLFAKGCFFRVIDGGMAYHQPPGKENETEREAGKSITLKIVKEK	413
Db	361	EEFNHGGEDVEFGYRLFAKGCFFRVIDGGMAYHQPPGKENETEREAGKSITLKIVKEK	420
Qy	414	VPYIYKLLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVVLEVCICNDGST	473
Db	421	VPYIYKLLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVVLEVCICNDGST	480
Qy	474	DNTLEVINLKYGNPRVRIMSKPNGGIASASNAVSAFAGYIIGQLSDSDYLEPDAVELC	533
Db	481	DNTLEVINLKYGNPRVRIMSKPNGGIASASNAVSAFAGYIIGQLSDSDYVLEPDAVELC	540
Qy	534	LKEFLDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL	593
Db	541	LKEFLDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL	600
Qy	594	TGDFENENIENAVDYMFLKSEVGFYHFKHNLKICYNRVLHGDNTSIKLGIOKKNHFVVVN	653
Db	601	TGDFENENIENAVDYMFLKSEVGFYHFKHNLKICYNRVLHGDNTSIKLGIOKKNHFVVVN	660
Qy	654	QSLNRQGINYNDKFDLDESRYIFNKTAEYQEEEDMLDKLQIONKDAKIAVSIFYP	713
Db	661	QSLNRQGINYNDKFDLDESRYIFNKTAEYQEEEDMLDKLQIONKDAKIAVSIFYP	720
Qy	714	NTNLGLVKLLNNIIEYNKNIFVILHLDKNHLPDIKKEILAFYHKKHQNILLNNDISYY	773
Db	721	NTNLGLVKLLNNIIEYNKNIFVILHLDKNHLPDIKKEILAFYHKKHQNILLNNDISYY	780
Qy	774	TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETH	833
Db	781	TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETH	840
Qy	834	DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQCS	893
Db	841	DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQCS	900
Qy	894	IDSVPYNTEDIWFQALLILEKKTGHVNKTSTLTTPWPERKLOWNEQIOSAKKGENI	953
Db	901	IDSVPYNTEDIWFQALLILEKKTGHVNKTSTLTTPWPERKLOWNEQIOSAKKGENI	960
Qy	954	PVNKFIINSITL 965	
Db	961	PVNKFIINSITL 972	
RESULT 9			
US-09-469-200-8			
; Sequence 8, Application US/09469200			
; GENERAL INFORMATION:			
; APPLICANT: Board of Regents of the University of Oklahoma			
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS			
; TITLE OF INVENTION: OF USE			
; FILE REFERENCE: 617022-7			

; CURRENT APPLICATION NUMBER: US/09/469,200
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 60/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida (PmCs)
US-09-469-200-8

Query Match 98.6%; Score 5017.5; DB 18; Length 972;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 953; Conservative 5; Mismatches 7; Indels 7; Gaps 1;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFQIIKCKELST-----NSYVS 53
Db 1 MNTLSQAIKAYNCNDYELALKLFKSAETGRKIVFQIIKCKELSTNSYVS 60
Qy 54 EDKNSVCDSSLDIATQLLSNVKLTLSSEKSLKNKWSITGKKSNAEIRKVELVP 113
Db 61 EDKNSVCDSSLDIATQLLSNVKLTLSSEKSLKNKWSITGKKSNAEIRKVELVP 120
Qy 114 KDFPKDLVLAHPDHVNDFTWYKNRKKSIGIKPVKNKIGLSIIPTENRSRILDTLACL 173
Db 121 KDFPKDLVLAHPDHVNDFTWYKNRKKSIGIKPVKNKIGLSIIPTENRSRILDTLACL 180
Qy 174 VNQNTNYPEVVVADGSGKENLTTVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQNTNYPEVVVADGSGKENLTTVQYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSLDCDMPAQQLWVHSYLTLELNDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSLDCDMPAQQLWVHSYLTLELNDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPSTISKGNISLDWRLHFKTDNLRCLDSPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 ETATNNPSTISKGNISLDWRLHFKTDNLRCLDSPRYFVAGNVAFSKWLNKVGWFD 360
Qy 354 EEFNWGGEDVEFGYRLFAKGFRRVIDGGMIAHOEPCKENETEREAGKSTLKVIREK 413
Db 361 EEFNWGGEDVEFGYRLFAKGFRRVIDGGMIAHOEPCKENETEREAGKSTLKVIREK 420
Qy 414 VPIYIRKLPTEDSHIHRIPLSYIIPAYNCANYTORCVDSALNQTVDLEVICNDGST 473
Db 421 VPIYIRKLPTEDSHIHRIPLSYIIPAYNCANYTORCVDSALNQTVDLEVICNDGST 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGSIASASNAVSFAKGYIGQLSDSDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGSIASASNAVSFAKGYIGQLSDSDYLEPDAVELC 540
Qy 534 LKEFLKDKTLACVYTTNRNVNDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTLACVYTTNRNVNDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
Qy 594 TDGFNENIENAVDYMFLKLSVGFKHLNKTICYNRVLHGDNTSKKGIQKKNHFPVVYN 653
Db 601 TDGFNENIENAVDYMFLKLSVGFKHLNKTICYNRVLHGDNTSKKGIQKKNHFPVVYN 660
Qy 654 QSLNRQGINYNKDFDLDDESKYIFNKTAQYQEMDMKDLKLTIONKDAKIAVSIFYP 713
Db 661 QSLNRQGINYNKDFDLDDESKYIFNKTAQYQEMDMKDLKLTIONKDAKIAVSIFYP 720
Qy 714 NTLNGLVKKLNIIIEYNKNFVIIHVDKNHLLTPDIKKEILAFYHKHVNILLNNDISY 773
Db 721 NTLNGLVKKLNIIIEYNKNFVIIHLDKNHLLTPDIKKEILAFYHKHVNILLNNDISY 780
Qy 774 TSNRLIKTEAHLNSNKLSQLNLCYEIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTH 833
Db 781 TSNRLIKTEAHLNSNKLSQLNLCYEIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTH 840

Qy 834 DWIEKINAHPPPKLLIKTYFNNDLRSMNVKGSQGMKMYALPHELLTIKIVITSQS 893
Db 841 DWIEKINAHPPPKLLIKTYFNNDLRSMNVKGSQGMKMYALPHELLTIKIVITSQS 900
Qy 894 IDSVPYENTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOMTNEQIOSAKKGENI 953
Db 901 IDSVPYENTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOMTNEQIOSAKKGENI 960
Qy 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972

RESULT 10
PCT-US99-07289-9
; Sequence 9, Application PC/TUS9907289
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 617022-7
; CURRENT APPLICATION NUMBER: PCT/US99/07289
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: 60/080,414
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 60/178,851
; EARLIER FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (17)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (62)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (90)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (113)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (158)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (164)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (192)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (201)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (206)
; OTHER INFORMATION: either Ile or Val
; FEATURE:

NAME/KEY: MOD_RES
LOCATION: (211)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (225)
OTHER INFORMATION: either Phe or Tyr
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (233)
OTHER INFORMATION: either Met or Leu
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (243)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (253)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (279)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (288)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (292)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (316)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (329)
OTHER INFORMATION: either Phe or Tyr
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (340)
OTHER INFORMATION: either Phe or Tyr
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (405)
OTHER INFORMATION: either Asn, Asp, Glu or Gln
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (439)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (744)
OTHER INFORMATION: either Ile or Val
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (952)
OTHER INFORMATION: either Asn, Asp, Glu or Gln

PCT-US99-07289-9

Query Match 90.4%; Score 4601.5; DB 1; Length 972;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 874; Conservative 28; Mismatches 63; Indels 7; Gaps 1;

QY 1 MNTLSQAIKAYNSNDYELAKLIFPKSAEYIGRKIVEFQIIKCKELSTNS-----YYS 53
Db 1 MNTLSQAIKAYNCNDYXALKLIFPKSAEYIGRKIVEFQIIKCKELSTNS-----YYS 60
QY 54 EDKKNVCSDDSDIATOLLISNVKKLTLSSEKNSLKNKWKSTITGKKSNAETIRKVELYP 113
Db 61 EXEKNVCSDDSDIATOLLISNVKKLTLSSEKNSLKNKWKSTITGKKSNAETIRKVELYP 120

QY 114 KDFPKDLVLAPLDPHVNDEFTWYNRKKSLGKTPVKNKNGLSIIITPTNRSRILDTIACL 173
Db 121 KDFPKDLVLAPLDPHVNDEFTWYNRKKRLGKTPKNGXGLSIIIXPTNRPAILDTIACL 180
QY 174 VNQNTNYPFEVVVADDSKENLLTIYOKYBQKLDIKVYRQKDYQYOLCAVNLGLRPAKY 233
Db 181 VNQNTNYPFEVVVADDSQEXLLPIYXOYEXKLDIRVQRKDNQXQACAARNXGLRLAKY 240
QY 234 DFVSLDCDMAPQOLWVHSYLTLELLEDNDIVLIGPRKYVDTHNTAEQFLNDPLYIESLP 293
Db 241 DFVSLDCDMAPQOLWVHSYLAELLEDDDTIIGPRKYXDTQNDIAEFLXASLIESLP 300
QY 294 ETATNNPSTISKGNISLDMRLEHFKTNDRLCDSPPRYFVAGNVAFSKWLAKVGHFD 353
Db 301 ETATNNPAAKGEGNXSLDMRLEQFETXNRLCDSPPRYFVAGNVAFSKWLAKVGHFD 360
QY 354 EEFNHWGGEDEVFGYRLFAKCFRVIDGGMAIHOEPGKNETEREAGKSTIKIVREK 413
Db 361 EEFNHWGGEDEVFGYRLFRKGCFFERTIDGGMAIHOEPGKNETEREAGKSTIKIVREK 420
QY 414 VPIYRKLLPTEDSHIRIPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGST 473
Db 421 VPIYRKLLPTEDSHINRPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCICNDGST 480
QY 474 DNTLEVINKLYGNPRVRIMSKPNGGSIASANAASVFAKGYIIGOLDSDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGSIASANAASVFAKGYIIGOLDSDDYLEPDAVELC 540
QY 534 LKEFLDKTCLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLDKTCLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 600
QY 594 TDGFENIENAVDYDMFLKSEVGKFKHLNLCYINRVLHGDNTSIKKLGIOKKNHFVVYN 653
Db 601 TDGFENIENAVDYDMFLKSEVGKFKHLNLCYINRVLHGDNTSIKKLGIOKKNHFVVYN 660
QY 654 QSLNRQGINYNKYDFDLDESRRKIFNKTAEYOEEMDKDLKLTIONKDAKIAVSIFYP 713
Db 661 QSLNRQGINYNKYDFDLDESRRKIFNKTAEYOEEMDKDLKLTIONKDAKIAVSIFYP 720
QY 714 NTLNGLVKKLNNIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISY 773
Db 721 NTLNGLVKKLNNIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHQNILLNNDISY 780
QY 774 TSNRLIKTEAHLNKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMNFSALTH 833
Db 781 TSNRLIKTEAHLNKLSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMNFSALTH 840
QY 834 DWIEKINAHPPPKLTKTYFENDNLRSMNVKASQGMFKYALPHELLTIIEKIVITSCOS 893
Db 841 DWIEKINAHPPPKLTKTYFENDNLRSMNVKASQGMFKYALPHELLTIIEKIVITSCOS 900
QY 894 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPMPWERKLTWNTQISAKKGENI 953
Db 901 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSTLTYPMPWERKLTWNTQISAKKGENI 960
QY 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

RESULT 11
US-09-283-402-9
; Sequence 9, Application US/09283402
; GENERAL INFORMATION:
; APPLICANT: Paul DeAngelis
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA
; TITLE OF INVENTION: MULTOCIDA AND USES THEREOF
; FILE REFERENCE: 5820.530
; CURRENT APPLICATION NUMBER: US/09/283,402
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; NAME/KEY: MOD_RES
; LOCATION: (17)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (62)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (90)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (113)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (158)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (164)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (201)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (206)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (211)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (225)
; OTHER INFORMATION: either Phe or Tyr
; NAME/KEY: MOD_RES
; LOCATION: (233)
; OTHER INFORMATION: either Met or Leu
; NAME/KEY: MOD_RES
; LOCATION: (243)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (253)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (279)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (288)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (292)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (316)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (329)
; OTHER INFORMATION: either Phe or Tyr
; NAME/KEY: MOD_RES
; LOCATION: (340)
; OTHER INFORMATION: either Phe or Tyr
; NAME/KEY: MOD_RES
; LOCATION: (405)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; NAME/KEY: MOD_RES
; LOCATION: (435)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
```

```
; LOCATION: (744)
; OTHER INFORMATION: either Ile or Val
; NAME/KEY: MOD_RES
; LOCATION: (952)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
US-09-283-402-9

Query Match          90.4%; Score 4601.5; DB 16; Length 972;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 874; Conservative 28; Mismatches 63; Indels 7; Gaps 1;

QY 1 MNTLSQAIKAYNSNDYELAKLFEKSAETVGRKIVFQIIEKKEKJSTNS-----YVS 53
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNTLSQAIKAYNCNDYXLAKLFEKSAETVGRKIVFQIIEKKEKJSTNSVNEANLSVN 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 54 EDKNSVCDSSLDIATOLLISNVKLTLSSEKNSLKNKWKSTGKSKSENAERKVELVP 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 EXEKNVCDSPLDIATOLLISNVKLTLSXSEKNSLKNKWKLTERRKSENAERAVALVP 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 114 KDFPKDLVLAPLDHVNDFTWYKNRKKSLGKPKVNKIGLSIIPFNRSRIIDITLACL 173
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 KDFPKDLVLAPLDHVNDFTWYKNRKKSLGKPKVNKIGLSIIPFNRSRIIDITLACL 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 174 VNOKTNPFEVYVADGSKENLTIIVQYEQKLDIYVRQKDYGYQLCAVRNGLRTAKY 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 VNOKTNPFEVYVADGSKENLTIIVQYEQKLDIYVRQKDYGYQLCAVRNGLRLAKY 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 234 DFVSLDCDMAPOQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEOLFNDPYLESLP 293
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 DFVSLDCDMAPOQLWVHSYLTLELLEDNDIVLIGPRKYVDTHNITAEOLFNDPYLESLP 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 294 ETATNNPSTSGNISLDWRLEHFKKTNLRLCDSPPRFYFVAGNVAFSEWLNKVGWFD 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ETATNNPSTSGNISLDWRLEHFKKTNLRLCDSPPRFYFVAGNVAFSEWLNKVGWFD 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 354 EEFNHWGGEDVEFGYRLFAKGCFFRVIDGMAIHOQPPGKENTETERAGKSIILKIVKEK 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 EEFNHWGGEDVEFGYRLFAKGCFFRVIDGMAIHOQPPGKENTETERAGKSIILKIVKEK 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 414 VPYIYRKLPIEDSHIRIPLVSIYIPAYNCANYIQRVDSALNQTVDLVLCVNDGST 473
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 VPYIYRKLPIEDSHIRIPLVSIYIPAYNCANYIQRVDSALNQTVDLVLCVNDGST 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 474 DNTLEVINKLYGNPRVRIMSKPNGGSIASNAASVAFKGYIIGQLSDDDYLEPDVYELC 533
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGSIASNAASVAFKGYIIGQLSDDDYLEPDVYELC 540
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHPFMTTIRAWHL 593
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHPFMTTIRAWHL 600
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 594 TDGFENENIENAVDYMFLKLSVGEKFKHLNKICYNRVLHGDNSTIKKLGIOKKNHFVVVN 653
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 TDGFENENIENAVDYMFLKLSVGEKFKHLNKICYNRVLHGDNSTIKKLGIOKKNHFVVVN 660
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 654 QSLNRQGINYNYDKFDDDESRYIFNKTAEYQEEEDMLKDLKIQNKDAKIAVSTFYP 713
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 QSLNRQGINYNYDKFDDDESRYIFNKTAEYQEEEDMLKDLKIQNKDAKIAVSTFYP 720
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 714 NTNLGLVKKLNIIEYNKNIIFVILHVDKNHLTPDKKEILAFYHKKHQVILLNNDISYY 773
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 NTNLGLVKKLNIIEYNKNIIFVILHVDKNHLTPDKKEILAFYHKKHQVILLNNDISYY 780
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 774 TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSIFVKNDYAYMKKYDVGNMFSALTH 833
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSIFVKNDYAYMKKYDVGNMFSALTH 840
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 834 DWIEKINAHPPFKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEIVITSCQS 893
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 DWIEKINAHPPFKKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEIVITSCQS 900
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 894 IDSVPEYNTEDIWFQFALLITILEKKTGHVFNKSTLTLYMPWPKLQNTNEOIQSAKGENI 953
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 901 IDSVPEYNTEDIWFALLILEKKTGHVFNKSTLTLYMPWPKLQWNEQIXSAKRGENI 960

QY 954 PVNKFINSITL 965
| | | | | | | | | |

Db 961 PVNKFINSITL 972

RESULT 12

US-09-469-200-9

Sequence 9, Application US/09469200

GENERAL INFORMATION:

APPLICANT: Board of Regents of the University of Oklahoma

TITLE OF INVENTION: NUCLEIC ACID ENCODING HYALURONAN SYNTHASE AND METHODS

TITLE OF INVENTION: OF USE

FILE REFERENCE: 617022-7

CURRENT APPLICATION NUMBER: US/09/469,200

CURRENT FILING DATE: 1999-12-21

PRIOR APPLICATION NUMBER: 60/080,414

PRIOR FILING DATE: 1998-04-02

PRIOR APPLICATION NUMBER: 60/178,851

PRIOR FILING DATE: 1998-10-26

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 972

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: consensus

NAME/KEY: MOD_RES

LOCATION: (17)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (62)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (90)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (113)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (158)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (164)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (192)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (201)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (206)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (211)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (225)

OTHER INFORMATION: either Phe or Tyr

NAME/KEY: MOD_RES

LOCATION: (233)

OTHER INFORMATION: either Met or Leu

NAME/KEY: MOD_RES

LOCATION: (243)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (253)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (479)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (288)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (292)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (316)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (329)

OTHER INFORMATION: either Phe or Tyr

NAME/KEY: MOD_RES

LOCATION: (340)

OTHER INFORMATION: either Phe or Tyr

NAME/KEY: MOD_RES

LOCATION: (405)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

NAME/KEY: MOD_RES

LOCATION: (439)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (744)

OTHER INFORMATION: either Ile or Val

NAME/KEY: MOD_RES

LOCATION: (952)

OTHER INFORMATION: either Asn, Asp, Glu or Gln

US-09-469-200-9

Query Match 90.4%; Score 4601.5; DB 18; Length 972;

Best Local Similarity 89.9%; Pred. No. 0;

Matches 874; Conservative 28; Mismatches 63; Indels 7; Gaps 1;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKVIYFEQIIEKKEKLSN-----YVS 53

Db 1 MNTLSQAIKAYNCNDYXLALKLFKSAETYGKVIYFEQIIEKKEKLSNPNVNEANLSVN 60

QY 54 EDKNSVCDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKKSSNAEIRKVELVP 113

Db 61 EXEKNVCDSPLDIATQLLSNVKLTLSSEKNSLKNKWSITGKKSSNAEIRKVELVP 120

QY 114 KDFPKDLVLAFLPDHVNDFTWYKNNKKSIGIKPVKNIGLSIIPTFNSRILDTLACL 173

Db 121 KDFPKDLVLAFLPDHVNDFTWYKNNKKSIGIKPVKNIGLSIIPTFNSRILDTLACL 180

QY 174 VNQKTNYPEVAVADDGSGENLTIIVQYEQKLDIKYVRQKDYGYOLCAVRNLGLTAKY 233

Db 181 VNQKTNYPEVAVADDGSGENLTIIVQYEQKLDIKYVRQKDYGYOLCAVRNLGLTAKY 240

QY 234 DFVSILDCDMPQQLWVHSYLTLELDDNDIVLIGPKRYVDTHNITAEQFLNDPYLIESLP 293

Db 241 DFVXILDCDMPQQLWVHSYLTLELDDNDIVLIGPKRYVDTHNITAEQFLNDPYLIESLP 300

QY 294 ETATNNPSITSGNISLDWRLEHFKKTDNRLCDSPPFRYFVAGNVAFSKENLKNYGVWF 353

Db 301 ETATNNPSITSGNISLDWRLEHFKKTDNRLCDSPPFRYFVAGNVAFSKENLKNYGVWF 360

QY 354 EEFNHWGGEDVEFGYRLFAKGCFFRVIDGMAIHQPPGKENTEREAGKSIYLVKIVKEK 413

Db 361 EEFNHWGGEDVEFGYRLFAKGCFFRVIDGMAIHQPPGKENTEREAGKSIYLVKIVKEK 420

QY 414 VPIYIEKLLPIEDSHTRIPVLSIYIPAYNCANYIQRVDSALNQVVDLEVCINDGST 473

Db 421 VPIYIEKLLPIEDSHTRIPVLSIYIPAYNCANYIQRVDSALNQVVDLEVCINDGST 480

QY 474 DNTLEVINKLYGNPRVRIMSKPENGGIASNAVAASFAGYIYIGQLDDDDYLEPDAVELC 533

Db 481 DNTLEVINKLYGNPRVRIMSKPENGGIASNAVAASFAGYIYIGQLDDDDYLEPDAVELC 540

QY 534 LKEFLKDKTLACVYTTNNRVNPDGSLIANGYNWPEFSREKLTMTAMTAHHEFMETIRAWHL 593

Db 541 LKEFLKDKTLACVYTTNNRVNPDGSLIANGYNWPEFSREKLTMTAMTAHHEFMETIRAWHL 600

QY 594 TDGFNENIENAVDYMFLKLVSEVGFKHLNLCYINRVLHGDNTSTIKKLGIOKKNHFVVVN 653
DB 601 TDGFNENIENAVDYMFLKLVSEVGFKHLNLCYINRVLHGDNTSTIKKLGIOKKNHFVVVN 660
QY 654 QSLNRQGINNYNDKFDLDDESRYIFNKTAQYQEMDMLDKLQIONKDAKIAVSIYFYP 713
DB 661 QSLNRQGINNYNDFDLDDESRYIFNKTAQYQEMDMLDKLQIONKDAKIAVSIYFYP 720
QY 714 NTLNGLVKKLNNIEYKNKNIIVLHVDKNHLPDPDIKEILAFYKHQVNNLLNNDISYY 773
DB 721 NTLNGLVKKLNNIEYKNKNIIVLHVDKNHLPDPDIKEILAFYKHQVNNLLNNDISYY 780
QY 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFYKNDISYAYMKYDVGMMFSALTH 833
DB 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFYKNDISYAYMKYDVGMMFSALTH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNLSRNMVKGASQGMFKYALPHELITIIKEVITSQCS 893
DB 841 DWIEKINAHPPFKLIKTYFNDNLSRNMVKGASQGMFKYALPHELITIIKEVITSQCS 900
QY 894 IDSVPYNTEDIWQFALLILEKKTGHVFNKSTLTLYMPWERKLTQWNEQIOSAKKGENI 953
DB 901 IDSVPYNTEDIWQFALLILEKKTGHVFNKSTLTLYMPWERKLTQWNEQIOSAKKGENI 960
QY 954 PVNKFIIINSITL 965
DB 961 PVNKFIIINSITL 972

RESULT 13
US-09-791-537-56855
; Sequence 56855, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56855
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-791-537-56855

Query Match 88.3%; Score 4493.5; DB 21; Length 972;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 844; Conservative 61; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAIKAYNSNDYELALKLFESAETGRKIVFOLIKCEKL----STNS---VVS 53
DB 1 MNTLSQAIKAYNSNDYELALKLFESAETGRKIVFOLIKCEKLKLSAHPVNSAHLNVN 60
QY 54 EDKNSVCDSSLDIATQALLSNVKKLTLSSEKSNLKNKWSITGKSENAEIRKVELVP 113
DB 61 KEKVNVCDSPLDIATQALLSNVKKLVLSDEKNTLKNKWLTKKSENAEVRVALVP 120
QY 114 KDFPKDLVLAFLPDHVNDFTYKNNRKSGLGKPVNKNKTGLSIIPTFNRSRILDTLACL 173
DB 121 KDFPKDLVLAFLPDHVNDFTYKNNRKSGLGKPVNKNKTGLSIIPTFNRSRILDTLACL 180
QY 174 VNOKTNPFEVYVADGSKENLLTVQKYEOKLDIKYVROKDYGYOLCAVRNLGLRTAKY 233
DB 181 VNOKTHYFEVIVTDGSGEDLSPIRQYENKLDIYVROKONGFOASARNWGLRLAKY 240
QY 234 DVFSTLDCDMPAQOLWVHSYLTLELNDIVLIGPRKYVDTHNITAEQFLNDPILIESLP 293
DB 241 DFIGLDCDMPANPLWVHSYVAELLEDLDTIIGPRKYIDTQIDPKDFLNNSLLESPL 300

QY 294 ETATNNPSITSGKNTSLDWRLEHFKKTDNLRCLDSFERYFVAGNVAFSEKWLKNKYGWFD 353
DB 301 EVKTHNSVAAGEGTVSLDWRLEQFEKTEENLRUSDSPFRFAAGNVAFSAKWLKNKSGFFD 360
QY 354 EEFNHWGGEDVEFGYRLFAGCGFFRVIDGGMATHQBPFGKENETEREAGKSITLKIVKEK 413
DB 361 EEFNHWGGEDVEFGYRLFAGCGFFRVIDGGMATHQBPFGKENETEREAGKSITLKIVKEK 420
QY 414 VPIYTRKLLPIEDSHIRPLVSIYIPAYNCANCIORCVDSALNQTVVLEVCICNDGST 473
DB 421 VPIYTRKLLPIEDSHIRPLVSIYIPAYNCANCIORCVDSALNQTVVLEVCICNDGST 480
QY 474 DNTLEVINKLNGNPNRVRIMSKPNGGSIASNAASVAFKGYIIGQLDSDDDYLPDPAVELC 533
DB 481 DNTLEVINKLNGNPNRVRIMSKPNGGSIASNAASVAFKGYIIGQLDSDDDYLPDPAVELC 540
QY 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHPMFTIRAWHL 593
DB 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHPMFTIRAWHL 600
QY 594 TDGFNENIENAVDYMFLKLVSEVGFKHLNLCYINRVLHGDNTSTIKKLGIOKKNHFVVVN 653
DB 601 TDGFNENIENAVDYMFLKLVSEVGFKHLNLCYINRVLHGDNTSTIKKLGIOKKNHFVVVN 660
QY 654 QSLNRQGINNYNDKFDLDDESRYIFNKTAQYQEMDMLDKLQIONKDAKIAVSIYFYP 713
DB 661 QSLNRQGINNYNDFDLDDESRYIFNKTAQYQEMDMLDKLQIONKDAKIAVSIYFYP 720
QY 714 NTLNGLVKKLNNIEYKNKNIIVLHVDKNHLPDPDIKEILAFYKHQVNNLLNNDISYY 773
DB 721 NTLNGLVKKLNNIEYKNKNIIVLHVDKNHLPDPDIKEILAFYKHQVNNLLNNDISYY 780
QY 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFYKNDISYAYMKYDVGMMFSALTH 833
DB 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFYKNDISYAYMKYDVGMMFSALTH 840
QY 834 DWIEKINAHPPFKLIKTYFNDNLSRNMVKGASQGMFKYALPHELITIIKEVITSQCS 893
DB 841 DWIEKINAHPPFKLIKTYFNDNLSRNMVKGASQGMFKYALPHELITIIKEVITSQCS 900
QY 894 IDSVPYNTEDIWQFALLILEKKTGHVFNKSTLTLYMPWERKLTQWNEQIOSAKKGENI 953
DB 901 IDSVPYNTEDIWQFALLILEKKTGHVFNKSTLTLYMPWERKLTQWNEQIOSAKKGENI 960
QY 954 PVNKFIIINSITL 965
DB 961 PVNKFIIINSITL 972

RESULT 14
US-10-011-768B-9
; Sequence 9, Application US/10011768B
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011.768B
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella Multocida
US-10-011-768B-9


```
Query Match      88.2%; Score 4490.5; DB 24; Length 972;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 843; Conservative 62; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAICAYNSNDYELALKEFSAETIGRKIVFEQIIKCKEKL-----STNS---YVS 53
DB 1 MNTLSQAICAYNSNDYQALALKEFSAEYIGRKIVFEQITCKEKLSAHPSVNSAHLNVN 60

QY 54 EDKNSVCDSSLDIATQALLSNVKKLTSESEKNSLKNKKSITGKSENAEIRKVELVP 113
DB 61 KEKVNVCDSPDIATQALLSNVKKLVSDSEKNTLANKWKLLEKKSSENAEVRVALVP 120

QY 114 KDPKDLVLAFLPDHVNDFWYKRRKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACL 173
DB 121 KDPKDLVLAFLPDHVNDFWYKRRKRLGKIPKPEHQHVGLSIIVTFNRPAILSLACL 180

QY 174 VNOKTNPFEVVDGSGKENLTIQVYEQKLDIKVVRQKDYQOLCAVRNGLRTAKY 233
DB 181 VNOKTHYFVIVTDDGSQEDLSPIIRQYENKLDIRYVRQKNGFQASARNMGLRLAKY 240

QY 234 DFVSLDCCDAPQOLWVHSYITELLEDDIVLIGPRKYVDTHNITAEQFLNDPVLIESLP 293
DB 241 DFVSLDCCDAPNPLWVHSYVAELLEDDITIIIGPRKYIDTQHDIPKDFLNNASLESPL 300

QY 294 ETATNNPISITSGNISLDMRLEHFKKTDNLRLCDSFPFYVAGNVAFSKEWLNKVGWFD 353
DB 301 EVKTNNSVAAKGEVTVSLOWRLEQFEKTEENLRSLSDSPFRFFAAGNVAFKAKWLNKSGFFD 360

QY 354 EEFNWGGEDVEFGYRLFPAKCCFFRVIDGMAIHQEPGPKENETEREAGKSITLKIYKEK 413
DB 361 EEFNWGGEDVEFGYRLFPAKCCFFRVIDGMAIHQEPGPKENETEREAGKSITLKIYKEK 420

QY 414 VPYIYRKLLPDESHIRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 473
DB 421 VPYIYRKLLPDESHIRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 480

QY 474 DNTLEVINKLYGNPRVIRMSKPNNGGSIASASNAAFSAKGYIIGOLDSDDDYLEDPAVELC 533
DB 481 DNTLEVINKLYGNPRVIRMSKPNNGGSIASASNAAFSAKGYIIGOLDSDDDYLEDPAVELC 540

QY 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPFESREKLTMTAMIAHFRMFTIRAWHL 593
DB 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPFESREKLTMTAMIAHFRMFTIRAWHL 600

QY 594 TDGFENIENAVDYDMFLKSEVGFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 653
DB 601 TDGFENIENAVDYDMFLKSEVGFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 660

QY 654 QSLNRQGINYNYDKFDDDESRYIFNKTAEYOEMDMKDLKLNQKDAKTAVSIFYP 713
DB 661 QSLNRQGINYNYDKFDDDESRYIFNKTAEYOEMDMKDLKLNQKDAKTAVSIFYP 720

QY 714 NTLNGLVKLNNIIEYNKNIIVILHVDKNHUTPDIKKEILAFYKHQVNNILNNDISY 773
DB 721 NTLNGLVKLNNIIEYNKNIIVILHVDKNHUTPDIKKEILAFYKHQVNNILNNDISY 780

QY 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMNFSALTH 833
DB 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKYDVGMNFSALTH 840

QY 834 DWIEKINAPPPKLIKTYFNDNDRSMNVKASQGMFKYALPHELLTIIKEVITSCOS 893
DB 841 DWIEKINAPPPKLIKTYFNDNDRSMNVKASQGMFKYALPHELLTIIKEVITSCOS 900

QY 894 IDSVPEYNTEDWFOFALLILEKKTGHVFNKSTLTYPMPWERKLTQWNEQIOSAKKGENI 953
DB 901 IDSVPEYNTEDWFOFALLILEKKTGHVFNKSTLTYPMPWERKLTQWNEQIOSAKKGENI 960

QY 954 PVNKFINSITL 965
DB 961 PVNKFINSITL 972
```

```
RESULT 15
US-10-011-771B-9
; Sequence 9, Application US/10011771B
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011/771B
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pastuerella Multocida
US-10-011-771B-9

Query Match      88.2%; Score 4490.5; DB 24; Length 972;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 843; Conservative 62; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAICAYNSNDYELALKEFSAETIGRKIVFEQIIKCKEKL-----STNS---YVS 53
DB 1 MNTLSQAICAYNSNDYQALALKEFSAEYIGRKIVFEQITCKEKLSAHPSVNSAHLNVN 60

QY 54 EDKNSVCDSSLDIATQALLSNVKKLTSESEKNSLKNKKSITGKSENAEIRKVELVP 113
DB 61 KEKVNVCDSPDIATQALLSNVKKLVSDSEKNTLANKWKLLEKKSSENAEVRVALVP 120

QY 114 KDPKDLVLAFLPDHVNDFWYKRRKSLGKIPVKNKIGLSIIPTFNRSRIIDITLACL 173
DB 121 KDPKDLVLAFLPDHVNDFWYKRRKRLGKIPKPEHQHVGLSIIVTFNRPAILSLACL 180

QY 174 VNOKTNPFEVVDGSGKENLTIQVYEQKLDIKVVRQKDYQOLCAVRNGLRTAKY 233
DB 181 VNOKTHYFVIVTDDGSQEDLSPIIRQYENKLDIRYVRQKNGFQASARNMGLRLAKY 240

QY 234 DFVSLDCCDAPQOLWVHSYITELLEDDIVLIGPRKYVDTHNITAEQFLNDPVLIESLP 293
DB 241 DFVSLDCCDAPNPLWVHSYVAELLEDDITIIIGPRKYIDTQHDIPKDFLNNASLESPL 300

QY 294 ETATNNPISITSGNISLDMRLEHFKKTDNLRLCDSFPFYVAGNVAFSKEWLNKVGWFD 353
DB 301 EVKTNNSVAAKGEVTVSLOWRLEQFEKTEENLRSLSDSPFRFFAAGNVAFKAKWLNKSGFFD 360

QY 354 EEFNWGGEDVEFGYRLFPAKCCFFRVIDGMAIHQEPGPKENETEREAGKSITLKIYKEK 413
DB 361 EEFNWGGEDVEFGYRLFPAKCCFFRVIDGMAIHQEPGPKENETEREAGKSITLKIYKEK 420

QY 414 VPYIYRKLLPDESHIRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 473
DB 421 VPYIYRKLLPDESHIRIPLVSIYIPAYNCANYIQRVDSALNQTVVDLEVCINCDSGT 480

QY 474 DNTLEVINKLYGNPRVIRMSKPNNGGSIASASNAAFSAKGYIIGOLDSDDDYLEDPAVELC 533
DB 481 DNTLEVINKLYGNPRVIRMSKPNNGGSIASASNAAFSAKGYIIGOLDSDDDYLEDPAVELC 540

QY 534 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPFESREKLTMTAMIAHFRMFTIRAWHL 593
DB 541 LKEFLKDKTLACVYTTNRNVNPDGSLIANGYNWPFESREKLTMTAMIAHFRMFTIRAWHL 600

QY 594 TDGFENIENAVDYDMFLKSEVGFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 653
DB 601 TDGFENIENAVDYDMFLKSEVGFKHLNKCICYNRVLHGDNTSIKLGIOKKNHFVVVN 660

QY 654 QSLNRQGINYNYDKFDDDESRYIFNKTAEYOEMDMKDLKLNQKDAKTAVSIFYP 713
DB 661 QSLNRQGINYNYDKFDDDESRYIFNKTAEYOEMDMKDLKLNQKDAKTAVSIFYP 720
```

Db 661 QSLNRQGTYYNYDEFDDDESRYIFNKTAEQEIDILKDIKTIQNKDAKIAVSIFYP 720
Qy 714 NTLGLVKKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHVNILLNNDISY 773
Db 721 NTLGLVKKLNNIIEYNKNIFVILHVDKNHLPDIIKKEILAFYHKHVNILLNNDISY 780
Qy 774 TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETH 833
Db 781 TSNRLIKTEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETH 840
Qy 834 DWIEKINAHPPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQS 893
Db 841 DWIEKINAHPPFKLIKTYFNDNDRSMNVKGASQGMFKYALPHELLTIIKEVITSCQS 900
Qy 894 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSTLTYPWPKLQWNEQIOSAKKGENI 953
Db 901 IDSVPEYNTEDIWFOFALLILEKKTGHVFNKTSTLTYPWPKLQWNEQIOSAKKGENI 960
Qy 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

Search completed: January 4, 2003, 02:36:14
Job time : 405 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:24:29 ; Search time 562 Seconds
(without alignments)
121.219 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGNIPIVKNKIINSITL 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 258925 seqs, 70596210 residues
Total number of hits satisfying chosen parameters: 258925

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5089	100.0	965	5	US-09-842-484A-2
2	5067	99.6	965	5	US-09-842-484A-4
3	5067	99.6	965	6	US-10-217-613-3
4	5067	99.6	965	6	US-10-217-613-8
5	4604.5	90.5	972	6	US-10-217-613-9
6	4490.5	88.2	972	5	US-09-469-200D-9
7	4486.5	88.2	972	6	US-10-309-560-8
8	3102.5	61.0	703	6	US-10-217-613-1
9	3102.5	61.0	703	6	US-10-217-613-7
10	2256.5	44.3	686	6	US-10-216-289-2
11	2256.5	44.3	686	6	US-10-216-289-4
12	333.5	6.6	842	5	US-09-134-000C-5749
13	299	5.9	331	6	US-10-264-213-152
14	262.5	5.2	534	5	US-09-134-000C-5087
15	248.5	4.9	330	5	US-09-134-000C-6392
16	240.5	4.7	1056	5	US-09-134-000C-5086
17	227	4.5	727	6	US-10-092-411A-4067
18	210	4.1	348	6	US-10-096-129-3
19	210	4.1	348	6	US-10-096-129-8
20	209	4.1	337	6	US-10-096-129-5
21	208.5	4.1	270	6	US-10-303-161-39
22	208.5	4.1	270	6	US-10-303-162-39
23	208.5	4.1	270	6	US-10-303-128-39
24	208.5	4.1	270	6	US-10-303-134-39
25	208.5	4.1	270	6	US-10-303-134-39
26	207	4.1	301	6	US-10-303-161-27

27	207	4.1	301	6	US-10-303-162-27	Sequence 27, Appl
28	207	4.1	301	6	US-10-303-118-27	Sequence 27, Appl
29	207	4.1	301	6	US-10-303-128-27	Sequence 27, Appl
30	207	4.1	301	6	US-10-303-134-27	Sequence 27, Appl
31	205	4.0	389	6	US-10-303-161-34	Sequence 34, Appl
32	205	4.0	389	6	US-10-303-162-34	Sequence 34, Appl
33	205	4.0	389	6	US-10-303-118-34	Sequence 34, Appl
34	205	4.0	389	6	US-10-303-128-34	Sequence 34, Appl
35	205	4.0	389	6	US-10-303-134-34	Sequence 34, Appl
36	204	4.0	651	6	US-10-142-143-6	Sequence 6, Appl
37	200.5	3.9	303	6	US-10-303-161-29	Sequence 29, Appl
38	200.5	3.9	303	6	US-10-303-162-29	Sequence 29, Appl
39	200.5	3.9	303	6	US-10-303-118-29	Sequence 29, Appl
40	200.5	3.9	303	6	US-10-303-128-29	Sequence 29, Appl
41	200.5	3.9	303	6	US-10-303-134-29	Sequence 29, Appl
42	200	3.9	617	6	US-10-142-143-2	Sequence 2, Appl
43	199	3.9	297	6	US-10-303-161-31	Sequence 31, Appl
44	199	3.9	297	6	US-10-303-162-31	Sequence 31, Appl
45	199	3.9	297	6	US-10-303-118-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-842-484A-2
; Sequence 2, Application US/09842484A
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAM
; FILE REFERENCE: 4605.003
; CURRENT APPLICATION NUMBER: US/09/842,484A
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-842-484A-2

Query Match 100.0%; Score 5089; DB 5; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVEFQIIKCKEKLSTNSVSDKKNV	60
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVEFQIIKCKEKLSTNSVSDKKNV	60
QY	61	CDSSLDIATQLLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVPKDFPKDL	120
Db	61	CDSSLDIATQLLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVPKDFPKDL	120
QY	121	VLAPLPHVNDFTWYKRNKSLGKIPVKNKIGLSIIPTFNRSRILDTITLACLVNQXTNY	180
Db	121	VLAPLPHVNDFTWYKRNKSLGKIPVKNKIGLSIIPTFNRSRILDTITLACLVNQXTNY	180
QY	181	PFVVVADGSGKENLLTIQVYEQKLDIKYVRQKDYQYQLCAVRNLGRTAKYDFVSILD	240
Db	181	PFVVVADGSGKENLLTIQVYEQKLDIKYVRQKDYQYQLCAVRNLGRTAKYDFVSILD	240
QY	241	CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNTAQFLNDPYLESLPETATNN	300
Db	241	CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNTAQFLNDPYLESLPETATNN	300
QY	301	PSITSGNLSLWRLHFKKTDNLRLCDSPFRYFVAGNVAFSKWLNKGVDFEENHWG	360
Db	301	PSITSGNLSLWRLHFKKTDNLRLCDSPFRYFVAGNVAFSKWLNKGVDFEENHWG	360
QY	361	GEDVEFGYRLFAKGCFFRVYIDGGMAIHOEPPGKENETEREAGKSITLKIYKVPYIYRK	420
Db	361	GEDVEFGYRLFAKGCFFRVYIDGGMAIHOEPPGKENETEREAGKSITLKIYKVPYIYRK	420

Db 361 GEDVEGYRLFAGCGFRVIDGGMATHOEPPGKENEERAGKSTILKIVKEKVPYIRK 420
QY 421 LLPEDSHIRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
Db 421 LLPEDSHIRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGASASNAVSAFAGYITGOLSDDDYLDPAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGASASNAVSAFAGYITGOLSDDDYLDPAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKNHVVVQSLNRQG 660
Db 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKNHVVVQSLNRQG 660
QY 661 INYNYDKFDLDESRYIFNKTAEYQOEEMDKDLKIQNKDAKIAVSIFYPNTLGLV 720
Db 661 INYNYDKFDLDESRYIFNKTAEYQOEEMDKDLKIQNKDAKIAVSIFYPNTLGLV 720
QY 721 KKLNNIIEYKNKIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
Db 721 KKLNNIIEYKNKIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTDHWEIKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTDHWEIKIN 840
QY 841 AHPPEFKLIKTYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCSIDSVPEY 900
Db 841 AHPPEFKLIKTYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCSIDSVPEY 900
QY 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENIPVNFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENIPVNFII 960
QY 961 NSITL 965
Db 961 NSITL 965

RESULT 2
US-09-842-484A-4
; Sequence 4, Application US/09842484A
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL L.
; TITLE OF INVENTION: CHONDROITIN SYNTHASE GENE AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 4605.003
; CURRENT APPLICATION NUMBER: US/09/842.484A
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ. ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ. ID NO 4
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-09-842-484A-4

Query Match 99.6%; Score 5067; DB 5; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MNTLSQAIKAYNSNDYELAKLIFERSAETYGRIKVEFQIIKCKELSTNSYSEDKKNSV 60
Db 1 MNTLSQAIKAYNSNDYELAKLIFERSAETYGRIKVEFQIIKCKELSTNSYSEDKKNSV 60
QY 61 CDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKKSNAERKVELYVPKDFPKDL 120
Db 61 CDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKKSNAERKVELYVPKDFPKDL 120

QY 121 VLAPLPDHVNDFTWYKNRKSLSGKIPVKNKNIIGLSIIIPTFNRSRIILDITLACLVNQKTN 180
Db 121 VLAPLPDHVNDFTWYKNRKSLSGKIPVKNKNIIGLSIIIPTFNRSRIILDITLACLVNQKTN 180
QY 181 PFEVVVADDSKKNLLTIVQYEQKLDIKYVRQKDYQYLCVARNLGLRTAKYDFVSILD 240
Db 181 PFEVVVADDSKKNLLTIVQYEQKLDIKYVRQKDYQYLCVARNLGLRTAKYDFVSILD 240
QY 241 CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTTAEOLFNDPYLESLPETATN 300
Db 241 CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTTAEOLFNDPYLESLPETATN 300
QY 301 PSITSKGNISLDWRLEHFKKTDNLRCLDSPERYFVAGNVAFSKWLKNKVGDFEENHWG 360
Db 301 PSITSKGNISLDWRLEHFKKTDNLRCLDSPERYFVAGNVAFSKWLKNKVGDFEENHWG 360
QY 361 GEDVEGYRLFAGCGFRVIDGGMATHOEPPGKENEERAGKSTILKIVKEKVPYIRK 420
Db 361 GEDVEGYRLFAGCGFRVIDGGMATHOEPPGKENEERAGKSTILKIVKEKVPYIRK 420
QY 421 LLPEDSHIRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
Db 421 LLPEDSHIRIPLVSIYIPAYNCANYIQCVDVSALNQTVVDLEVCICNDGSDNTLEVI 480
QY 481 NKLYGNPRVRIMSKPNGGASASNAVSAFAGYITGOLSDDDYLDPAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGASASNAVSAFAGYITGOLSDDDYLDPAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNVPDGLSIANGYNWPEFSREKLTITAMIAHFRMFTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKNHVVVQSLNRQG 660
Db 601 IENAVDYDMFLKSEVGKFKHLNKCYNRVLHGDNTSIKKLGIOQKNHVVVQSLNRQG 660
QY 661 INYNYDKFDLDESRYIFNKTAEYQOEEMDKDLKIQNKDAKIAVSIFYPNTLGLV 720
Db 661 INYNYDKFDLDESRYIFNKTAEYQOEEMDKDLKIQNKDAKIAVSIFYPNTLGLV 720
QY 721 KKLNNIIEYKNKIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
Db 721 KKLNNIIEYKNKIFVILHVDKNHLPDIKKEILAFYHKHQNVLNNDISYYSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTDHWEIKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTDHWEIKIN 840
QY 841 AHPPEFKLIKTYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCSIDSVPEY 900
Db 841 AHPPEFKLIKTYFNDNDRSMNVKASOGMPKMYALPHELLTIIKEVITSCSIDSVPEY 900
QY 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENIPVNFII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKSTLTLYMPWERKLOWTNEQIOSAKKGENIPVNFII 960
QY 961 NSITL 965
Db 961 NSITL 965

RESULT 3
US-10-217-613-3
; Sequence 3, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA ANI
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 35341.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402

```
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-217-613-3

Query Match          99.6%; Score 5067; DB 6; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
DB 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
QY 61 CDSSLDIATOLLNSVKKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVPKDPKDL 120
DB 61 CDSSLDIATOLLNSVKKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVPKDPKDL 120
QY 121 VLAPLPHVNDFTWYKNNRKSGLGIRPVKNKNGLSIIPTFNRSRIIDITLACLVNQKTN 180
DB 121 VLAPLPHVNDFTWYKNNRKSGLGIRPVKNKNGLSIIPTFNRSRIIDITLACLVNQKTN 180
QY 181 PFEVVADGSGKENLLTIIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
DB 181 PFEVVADGSGKENLLTIIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATN 300
DB 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATN 300
QY 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKWLNKVGWDFDEPNHWG 360
DB 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKWLNKVGWDFDEPNHWG 360
QY 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKNETEREAGKSITLKIVKEKVPYIYRK 420
DB 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKNETEREAGKSITLKIVKEKVPYIYRK 420
QY 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDLEVCIKNDGSTDNTLEVI 480
DB 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDLEVCIKNDGSTDNTLEVI 480
QY 481 NKLYGNPNRVRIMSKPNGGSIASNAVSAFAGYIYGQDSDDDYLPDPAVELCLKEFLKD 540
DB 481 NKLYGNPNRVRIMSKPNGGSIASNAVSAFAGYIYGQDSDDDYLPDPAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMTAHFRMPTIRAWHLTDGFNEN 600
DB 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMTAHFRMPTIRAWHLTDGFNEN 600
QY 601 IENAVDYDMFLKSLSEVGKFKHLNKICYNRVLHGDNTSIKKGITQKNHFFVWVQSLNRQ 660
DB 601 IENAVDYDMFLKSLSEVGKFKHLNKICYNRVLHGDNTSIKKGITQKNHFFVWVQSLNRQ 660
QY 661 INYNYDKPDDJDESRYIFNKTAEQEEMDLKDLKLQNKDQAKTAVSIFYPNTLNGVLV 720
DB 661 INYNYDKPDDJDESRYIFNKTAEQEEMDLKDLKLQNKDQAKTAVSIFYPNTLNGVLV 720
QY 721 KKLNNIIEYNKNFVLIILHVDKNHLTPDIKKEILAFYHKHQVNILLNNDISYVTSNRLIK 780
DB 721 KKLNNIIEYNKNFVLIILHVDKNHLTPDIKKEILAFYHKHQVNILLNNDISYVTSNRLIK 780
QY 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVNDSYAYMKKVDGMNFSALTHDWTIEKIN 840
DB 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVNDSYAYMKKVDGMNFSALTHDWTIEKIN 840
```

```
QY 841 AHPPEKKLIKTYFNDNDLRSMNVKASQGMFKYALPHELLTIIKEVITSQSIDSVPEY 900
DB 841 AHPPEKKLIKTYFNDNDLRSMNVKASQGMFKYALPHELLTIIKEVITSQSIDSVPEY 900
QY 901 NTEDIWFOFALLTLEKKTGHVFNKSTLTLYMPWERKLQWNTNEQIQSAKKGENIPVKNFII 960
DB 901 NTEDIWFOFALLTLEKKTGHVFNKSTLTLYMPWERKLQWNTNEQIQSAKKGENIPVKNFII 960
QY 961 NSITL 965
DB 961 NSITL 965

RESULT 4
US-10-217-613-8
; Sequence 8, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AT
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-217-613-8

Query Match          99.6%; Score 5067; DB 6; Length 965;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 960; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
DB 1 MNTLSQAIKAYNSNDYELALKLFKSAETGRKIVFEQIIKCKEKLSTNSYVSEDKKNV 60
QY 61 CDSSLDIATOLLNSVKKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVPKDPKDL 120
DB 61 CDSSLDIATOLLNSVKKLTLSSEKNSLKNKWSITGKKSNAEIRKVELVPKDPKDL 120
QY 121 VLAPLPHVNDFTWYKNNRKSGLGIRPVKNKNGLSIIPTFNRSRIIDITLACLVNQKTN 180
DB 121 VLAPLPHVNDFTWYKNNRKSGLGIRPVKNKNGLSIIPTFNRSRIIDITLACLVNQKTN 180
QY 181 PFEVVADGSGKENLLTIIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
DB 181 PFEVVADGSGKENLLTIIVQKYEQKLDIKYVRQKDYGYQLCAVRNLGLRTAKYDFVSILD 240
QY 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATN 300
DB 241 CDMAPOOLWVHSLYTELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESPETATN 300
QY 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKWLNKVGWDFDEPNHWG 360
DB 301 PSITSGKNSLDWRLEHFKKTDNLRCLDSPFRYFVAGNVAFSKWLNKVGWDFDEPNHWG 360
QY 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKNETEREAGKSITLKIVKEKVPYIYRK 420
DB 361 GEDVEFGYRLFAKGCFFRVIDGGMAYHQEPGKNETEREAGKSITLKIVKEKVPYIYRK 420
QY 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDLEVCIKNDGSTDNTLEVI 480
DB 421 LLPIEDSHIHRPLVSIYIPAYNCANYIQRCVDSALNQTVDLEVCIKNDGSTDNTLEVI 480
```

```
QY 481 NKLYGNPRVRIMSKPNGGIASASNAVSPAKGYIIGQLDDDDYLEPDVAVELCLKEFLKD 540
Db 481 NKLYGNPRVRIMSKPNGGIASASNAVSPAKGYIIGQLDDDDYLEPDVAVELCLKEFLKD 540
QY 541 KTLACVYTTNRNPNPDSGLIANGYNWPEFSREKLTAMIAHHFMTTIRAWHLTDGFEN 600
Db 541 KTLACVYTTNRNPNPDSGLIANGYNWPEFSREKLTAMIAHHFMTTIRAWHLTDGFEN 600
QY 601 IENAVDYDMFLKLSEVGFKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVVQSLNRQ 660
Db 601 IENAVDYDMFLKLSEVGFKFKHLNKCYNRVLHGDNTSIKKLGIOKKNHFFVVVQSLNRQ 660
QY 661 INYNYDKFDLDESRYIENKTAEQEEMDLKDLKLNQKDAKIAVSIFYPNTLNGLV 720
Db 661 INYNYDKFDLDESRYIENKTAEQEEMDLKDLKLNQKDAKIAVSIFYPNTLNGLV 720
QY 721 KKLNNIIEYKNTFVILHLVDKKNLTPDIKKEILAFYKHQVNLNNDISYYTTSNRLIK 780
Db 721 KKLNNIIEYKNTFVILHLVDKKNLTPDIKKEILAFYKHQVNLNNDISYYTTSNRLIK 780
QY 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWIEKIN 840
Db 781 TEAHLNINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSAETHDWIEKIN 840
QY 841 AHPPEFKLIKTYFNDNDRSNVKGASQGMFKYALPHELLTIIKEVITSCQSIDSVPEY 900
Db 841 AHPPEFKLIKTYFNDNDRSNVKGASQGMFKYALPHELLTIIKEVITSCQSIDSVPEY 900
QY 901 NTEIDWQFALLILEKKTGHVFNKTSITLYMPWERKLOWNEQIQSAKKGENIPVNFII 960
Db 901 NTEIDWQFALLILEKKTGHVFNKTSITLYMPWERKLOWNEQIQSAKKGENIPVNFII 960
QY 961 NSITL 965
Db 961 NSITL 965

RESULT 5
US-10-217-613-9
; Sequence 9, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AND US
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (17)..(17)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (62)..(62)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (90)..(90)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (113)..(113)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (158)..(158)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (164)..(164)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (192)..(192)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (201)..(201)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (206)..(206)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (211)..(211)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (225)..(225)
; OTHER INFORMATION: either Phe or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (233)..(233)
; OTHER INFORMATION: either Met or Leu
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (243)..(243)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (253)..(253)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (279)..(279)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (288)..(288)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (292)..(292)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (316)..(316)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (329)..(329)
; OTHER INFORMATION: either Phe or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (340)..(340)
; OTHER INFORMATION: either Phe or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (405)..(405)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
; FEATURE:
```



```
; NAME/KEY: MISC_FEATURE
; LOCATION: (439)..(439)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (744)..(744)
; OTHER INFORMATION: either Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (952)..(952)
; OTHER INFORMATION: either Asn, Asp, Glu or Gln
US-10-217-613-9

Query Match      90.5%; Score 4604.5; DB 6; Length 972;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 874; Conservative 29; Mismatches 62; Indels 7; Gaps 1;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIICKKEKLSN-----YVS 53
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIICKKEKLSN-----YVS 60
Qy 54 EDKNSVCDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKSENAEIRKVELVP 113
Db 61 EXEKNVCDSPDLATQLLSNVKLTLSXSEKNSLKNKWLITEKKSENAEIRKVELVP 120
Qy 114 KDFPKDLVLAIPHNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSRIIDITLACL 173
Db 121 KDFPKDLVLAIPHNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSRIIDITLACL 180
Qy 174 VNQKTYPEVVVADGSGKENLTTIVQYEQKLDITKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTYPEVVVADGSGQEXLLPLTXROYEXKLDITRYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSTLDCDMPAQLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSTLDCDMPAQLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPNSITSGNISDLWLEHFKTDNLRCLDPSERYFVAGNVAFSKENLKVGFWD 353
Db 301 ETATNNPNSITSGNISDLWLEHFKTDNLRCLDPSERYFVAGNVAFSKENLKVGFWD 360
Qy 354 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 413
Db 361 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 420
Qy 414 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCINDGST 473
Db 421 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCINDGST 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGIIASASNAVSAFAGYVIGLSDSDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVRIMSKPNGGIIASASNAVSAFAGYVIGLSDSDYLEPDAVELC 540
Qy 534 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
Qy 594 TDGFENIENAVDMDFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIQKKNHFVVYN 653
Db 601 TDGFENIENAVDMDFLKSEVGKFKHLNKICYNRVLHGDNTSIKKLGIQKKNHFVVYN 660
Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAQYQEMDMLKDLKLTQNKDKAIAVSIFYP 713
Db 661 QSLNRQGINNYNDKFDLDESRYIFNKTAQYQEMDMLKDLKLTQNKDKAIAVSIFYP 720
Qy 714 NTLNGLVKKLNNIIEYNKNIFVILHVDKNHLTPDKKEILAFYHKHVNILLNNDISY 773
Db 721 NTLNGLVKKLNNIIEYNKNIFVILHVDKNHLTPDKKEILAFYHKHVNILLNNDISY 780
Qy 774 TSNRLIKTEAHLNLSNKLNSOLNCEYIIFDNHDSLFVKNDISYAYKKYDVGMNFSALTH 833
Db 781 TSNRLIKTEAHLNLSNKLNSOLNCEYIIFDNHDSLFVKNDISYAYKKYDVGMNFSALTH 840
```

```
Qy 834 DWIEKINAHPPFKKLIKTYFENDNDRSMNVKGASQGMFKMYALPHELLTIKEIVITSCQS 893
Db 841 DWIEKINAHPPFKKLIKTYFENDNDRSMNVKGASQGMFKMYALPHELLTIKEIVITSCQS 900
Qy 894 IDSVPEYNTEDIWFQPALILILEKKTCHGVFNKNTSTLTLYMPWERKLTQWNTBOIQSAKGENI 953
Db 901 IDSVPEYNTEDIWFQPALILILEKKTCHGVFNKNTSTLTLYMPWERKLTQWNTBOIQSAKGENI 960
Qy 954 PVNKKFIINSITL 965
Db 961 PVNKKFIINSITL 972

RESULT 6
US-09-469-200D-9
; Sequence 9, Application US/09469200D
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/09/469,200D
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pastuerella Multocida
US-09-469-200D-9

Query Match      88.2%; Score 4490.5; DB 5; Length 972;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 843; Conservative 62; Mismatches 60; Indels 7; Gaps 2;

Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIICKKEKLSN-----YVS 53
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVFQIICKKEKLSN-----YVS 60
Qy 54 EDKNSVCDSSLDIATQLLSNVKLTLSSEKNSLKNKWSITGKSENAEIRKVELVP 113
Db 61 EXEKNVCDSPDLATQLLSNVKLTLSXSEKNSLKNKWLITEKKSENAEIRKVELVP 120
Qy 114 KDFPKDLVLAIPHNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSRIIDITLACL 173
Db 121 KDFPKDLVLAIPHNDFTWYKNRKSLSGKPKVKNKIGLSIIPTFNRSRIIDITLACL 180
Qy 174 VNQKTYPEVVVADGSGKENLTTIVQYEQKLDITKYVRQKDYGYQLCAVRNLGLRTAKY 233
Db 181 VNQKTYPEVVVADGSGQEXLLPLTXROYEXKLDITRYVRQKDYGYQLCAVRNLGLRTAKY 240
Qy 234 DFVSTLDCDMPAQLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 293
Db 241 DFVSTLDCDMPAQLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLIESLP 300
Qy 294 ETATNNPNSITSGNISDLWLEHFKTDNLRCLDPSERYFVAGNVAFSKENLKVGFWD 353
Db 301 ETATNNPNSITSGNISDLWLEHFKTDNLRCLDPSERYFVAGNVAFSKENLKVGFWD 360
Qy 354 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 413
Db 361 EEFNHGGEDEVFGYRLFAKGCFFRVIDGMAIHQEPGKENETREAGKSTTLKIVREK 420
Qy 414 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCINDGST 473
Db 421 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRVDSALNQTVDVLEVCINDGST 480
Qy 474 DNTLEVINKLYGNPRVRIMSKPNGGIIASASNAVSAFAGYVIGLSDSDYLEPDAVELC 533
```

```
Db 481 DNTLEVINKLYGNNPRVIRMSKPNNGIASASNAAVSFAGYYIGQLSDDDYLEPDAVELC 540
QY 534 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFENIENAVDYDMFLKSEVKGKHLNKCYNRVLHGDNTSIKKLGIOKKNHFVVVN 653
Db 601 TDGFENIENAVDYDMFLKSEVKGKHLNKCYNRVLHGDNTSIKKLGIOKKNHFVVVN 660
QY 654 QSLNRQGINVYNDKFDLDESRYIFNKTAEOEEMDLKDLKIONKDAKIAVSIFYP 713
Db 661 QSLNRQGINVYNDKFDLDESRYIFNKTAEOEEMDLKDLKIONKDAKIAVSIFYP 720
QY 714 NTLNGLVKLLNIIETYNKNTFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 773
Db 721 NTLNGLVKLLNIIETYNKNTFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 780
QY 774 TSRLIKTEAHLNINKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWNFSAETH 833
Db 781 TSRLIKTEAHLNINKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWNFSAETH 840
QY 834 DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQOS 893
Db 841 DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQOS 900
QY 894 IDSVPYNTEDIWFOFALLILEKKTGHVFNKTSLTLYMPWERKLOWTNEQIOAKKGENI 953
Db 901 IDSVPYNTEDIWFOFALLILEKKTGHVFNKTSLTLYMPWERKLOWTNEQIOAKKGENI 960
QY 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

RESULT 7
US-10-309-560-8
; Sequence 8, Application US/10309560
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMAR, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-309-560-8

Query Match 88.2%; Score 4486.5; DB 6; Length 972;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQATKAYNSNDYELALKLFKSAETYGKRIYVEFOIICKEKL----STNS---YVS 53
Db 1 MNTLSQATKAYNSNDYELALKLFKSAETYGKRIYVEFOIICKEKL---STNS---YVS 60
QY 54 EDKKNVCSDDLDTATQLLSNVKKTLSSEKNSLKNKKSITGKKSSENAEIRKVELVP 113
Db 61 KEKVNVCSDDLDTATQLLSNVKKTLSSEKNSLKNKKSITGKKSSENAEIRKVELVP 120
QY 114 KDFPKDLVLPDHDVNTFWYKKNKSLGKPVKNKNGLSIIITFFNRSRILDTITLACL 173
Db 121 KDFPKDLVLPDHDVNTFWYKKNKSLGKPVKNKNGLSIIITFFNRSRILDTITLACL 180
QY 174 VNQKTNYPFEVVAVDDGSKENLLTIVQKYEOKLDTIKYVRQKDYQOLCAVRNIGLRTAKY 233
```

```
Db 181 VNQKTHYPFEVITVDGSGQEDLSPIIQRYENKLDIRVROKNDGFOASAARNMGLRLAKY 240
QY 234 DFTVSLDCDMAPOOLWVHSYLTLELLENDIVLIGPRKYVDTHNTITAEQFLNDPVLIESLP 293
Db 241 DFTGLDCDMAPNPLWVHSYVAELLEDDDLTIIGPRKYIDTQHIIDPKDFLNNASLLESPL 300
QY 294 ETATNNPSTTSKGNISLDWRLEHFKKTDNLRLCDSPFRVFNAGNVAFAFSEKWLAKVGFWD 353
Db 301 EVKTNNSVAAKGEGTSLDWRLEQFEXTENLRLSDSFPRFAAGNVAFAFSEKWLAKVGFWD 360
QY 354 EEFNHWGEGDEVEFGYRLFAKCGCFRVIDGGMAIHQEPGPKENETERAGKSIITIKVKEK 413
Db 361 EEFNHWGEGDEVEFGYRLFRYGSFETIDGIMAYHQEPGPKENETERAGKSIITIDIMREK 420
QY 414 VPIYRKLLPIEDSHIRIPLVSIYIPAYNCANYIQRCVDSALNQTVVVLEVCINDGST 473
Db 421 VPIYRKLLPIEDSHIRIPLVSIYIPAYNCANYIQRCVDSALNQTVVVLEVCINDGST 480
QY 474 DNTLEVINKLYGNNPRVIRMSKPNNGIASASNAAVSFAGYYIGQLSDDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNNPRVIRMSKPNNGIASASNAAVSFAGYYIGQLSDDDYLEPDAVELC 540
QY 534 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLDKDTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMFTIRAWHL 600
QY 594 TDGFENIENAVDYDMFLKSEVKGKHLNKCYNRVLHGDNTSIKKLGIOKKNHFVVVN 653
Db 601 TDGFENIENAVDYDMFLKSEVKGKHLNKCYNRVLHGDNTSIKKLGIOKKNHFVVVN 660
QY 654 QSLNRQGINVYNDKFDLDESRYIFNKTAEOEEMDLKDLKIONKDAKIAVSIFYP 713
Db 661 QSLNRQGINVYNDKFDLDESRYIFNKTAEOEEMDLKDLKIONKDAKIAVSIFYP 720
QY 714 NTLNGLVKLLNIIETYNKNTFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 773
Db 721 NTLNGLVKLLNIIETYNKNTFVILHVDKNHLPDIKKEILAFYHKHGVNILLNNDISYY 780
QY 774 TSRLIKTEAHLNINKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWNFSAETH 833
Db 781 TSRLIKTEAHLNINKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWNFSAETH 840
QY 834 DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQOS 893
Db 841 DWIEKINAHPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIIKEVITSQOS 900
QY 894 IDSVPYNTEDIWFOFALLILEKKTGHVFNKTSLTLYMPWERKLOWTNEQIOAKKGENI 953
Db 901 IDSVPYNTEDIWFOFALLILEKKTGHVFNKTSLTLYMPWERKLOWTNEQIOAKKGENI 960
QY 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972

RESULT 8
US-10-217-613-1
; Sequence 1, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AND
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-217-613-1

Query Match 61.0%; Score 3102.5; DB 6; Length 703;
Best Local Similarity 82.6%; Pred. No. 1.5e-287;
Matches 581; Conservative 57; Mismatches 58; Indels 7; Gaps 2;

Qy 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIKVEFOIICKEKL-----STNS---YVS 53
Db 1 MNTLSQAIKAYNSNDYQALALKLFPEKSAEYIGRKIVFEQITCKEKLUSAHPVSNVSAHLSVN 60
Qy 54 EDKNSVCSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKKSENAEIRKVELVP 113
Db 61 KEKVNVCDSPDLIATQALLSNVKKLVLSDEKNTLKNKKLTKKSENAEYRAVALVP 120
Qy 114 KDFPKDLVLAIPHVDNFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSRILDTITLACL 173
Db 121 KDFPKDLVLAIPHVDNFTWYKRRKRLGKIPKHQHVGLSIIPTFNRPAISITLACL 180
Qy 174 VNQKNTPEVVDAGDSKENLLTIVQYEOKLDIKYVRQKDYGYQLCAVRNGLRTAKY 233
Db 181 VNQKTHYPEVIVTDGSDQESLPIIROYENKLDIYVRQKDNQFQASAAARMGLRLAKY 240
Qy 234 DFVSTLDCDMPQOVLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLP 293
Db 241 DFIGLLDCDMPNPLWVHSYVAELLEDLTIIGPRKYIDTQHDIPKDFLNNASLESLEP 300
Qy 294 ETATNNPITSKGNISLDWRLEHFKKTDNRLCDSPPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 EVKTNNSVAAKGEVTVSLDWRLEQFEKTENLRSLDSPFRFFAAGNVAFKAKWLNKSGFFD 360
Qy 354 EEFNHMGDEDFVGYRLFAKCGFFRIDGMAIHQEPGKNETEREAGKSIITLKIVKEK 413
Db 361 EEFNHMGDEDFVGYRLFRYGSFEKTDIGIMAYHQEPGKNETEREAGKNTLIDIMREK 420
Qy 414 VPYIRKLLPTEDSHRIPLVSIYPAYNCANYTORCVDSALNQTVDLEVCICNDGST 473
Db 421 VPYIRKLLPTEDSHINRVLVSIYPAYNCANYIQRVDSALNQTVDLEVCICNDGST 480
Qy 474 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAKGYIIGLSDDDYLEPDAVELC 533
Db 541 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 600
Qy 594 TDGFNEKIENAVDYMFLKSEVGKFKHLNKICYNRVLRHGDNTSIKKLGIOKKNHFVVVN 653
Db 601 TDGFNEKIENAVDYMFLKSEVGKFKHLNKICYNRVLRHGDNTSIKKLGIOKKNHFVVVN 660
Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAEYQEBMDMLKDL 696
Db 661 QSLNRQGIITYNYDEDFDLDESRYIFNKTAEYQEEIDILKDI 703

RESULT 9
US-10-217-613-7
; Sequence 7, Application US/10217613
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTURELLA MULTOCIDA AND US
; FILE REFERENCE: 35541.081
; CURRENT APPLICATION NUMBER: US/10/217,613
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414

; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-217-613-7

Query Match 61.0%; Score 3102.5; DB 6; Length 703;
Best Local Similarity 82.6%; Pred. No. 1.5e-287;
Matches 581; Conservative 57; Mismatches 58; Indels 7; Gaps 2;

Qy 1 MNTLSQAIKAYNSNDYELALKLFPEKSAETYGKRIKVEFOIICKEKL-----STNS---YVS 53
Db 1 MNTLSQAIKAYNSNDYQALALKLFPEKSAEYIGRKIVFEQITCKEKLUSAHPVSNVSAHLSVN 60
Qy 54 EDKNSVCSSLDIATQALLSNVKKLTLSSEKNSLKNKWSITGKKSENAEIRKVELVP 113
Db 61 KEKVNVCDSPDLIATQALLSNVKKLVLSDEKNTLKNKKLTKKSENAEYRAVALVP 120
Qy 114 KDFPKDLVLAIPHVDNFTWYKRRKSLGKIPVKNKIGLSIIPTFNRSRILDTITLACL 173
Db 121 KDFPKDLVLAIPHVDNFTWYKRRKRLGKIPKHQHVGLSIIPTFNRPAISITLACL 180
Qy 174 VNQKNTPEVVDAGDSKENLLTIVQYEOKLDIKYVRQKDYGYQLCAVRNGLRTAKY 233
Db 181 VNQKTHYPEVIVTDGSDQESLPIIROYENKLDIYVRQKDNQFQASAAARMGLRLAKY 240
Qy 234 DFVSTLDCDMPQOVLWVHSYLTLELDDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLP 293
Db 241 DFIGLLDCDMPNPLWVHSYVAELLEDLTIIGPRKYIDTQHDIPKDFLNNASLESLEP 300
Qy 294 ETATNNPITSKGNISLDWRLEHFKKTDNRLCDSPPRYFVAGNVAFSKWLNKVGWFD 353
Db 301 EVKTNNSVAAKGEVTVSLDWRLEQFEKTENLRSLDSPFRFFAAGNVAFKAKWLNKSGFFD 360
Qy 354 EEFNHMGDEDFVGYRLFAKCGFFRIDGMAIHQEPGKNETEREAGKSIITLKIVKEK 413
Db 361 EEFNHMGDEDFVGYRLFRYGSFEKTDIGIMAYHQEPGKNETEREAGKNTLIDIMREK 420
Qy 414 VPYIRKLLPTEDSHRIPLVSIYPAYNCANYIQRVDSALNQTVDLEVCICNDGST 473
Db 421 VPYIRKLLPTEDSHINRVLVSIYPAYNCANYIQRVDSALNQTVDLEVCICNDGST 480
Qy 474 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAKGYIIGLSDDDYLEPDAVELC 533
Db 481 DNTLEVINKLYGNPRVIMSKPNGGSIASNAAVSFAKGYIIGLSDDDYLEPDAVELC 540
Qy 534 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 593
Db 541 LKEFLKDKTLACVYTTNRVNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL 600
Qy 594 TDGFNEKIENAVDYMFLKSEVGKFKHLNKICYNRVLRHGDNTSIKKLGIOKKNHFVVVN 653
Db 601 TDGFNEKIENAVDYMFLKSEVGKFKHLNKICYNRVLRHGDNTSIKKLGIOKKNHFVVVN 660
Qy 654 QSLNRQGINNYNDKFDLDESRYIFNKTAEYQEBMDMLKDL 696
Db 661 QSLNRQGIITYNYDEDFDLDESRYIFNKTAEYQEEIDILKDI 703

RESULT 10
US-10-216-289-2
; Sequence 2, Application US/10216289
; GENERAL INFORMATION:
; APPLICANT: NINOMIYA, TOSHIO
; APPLICANT: SUMIURA, NOBUO
; APPLICANT: KIMATA, KOJI
; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME
; FILE REFERENCE: 226882USO

; CURRENT APPLICATION NUMBER: US/10/216,289
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: JP 2001-244685
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: JP 2001-324127
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: JP 2002-103136
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-216-289-2

Query Match 44.3%; Score 2256.5; DB 6; Length 686;
Best Local Similarity 61.3%; Pred. No. 9.8e-207;
Matches 419; Conservative 112; Mismatches 148; Indels 5; Gaps 3;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
Db 1 MSTLNQAINLYKNKNYQALSLFEKVAEIVDSVSWVEANIKLCOTALNLSEVDKLNKAV 60
QY 61 CDSLSDIATQLLLSNVKKLTLSSEKSNLKNWKSITGKKSNAEIRKVELVPKDFPKDL 120
Db 61 ID--IDAATKIMCSNAKAIISLNEVEKNEIISKYREITAKKSERAEELKEVEPIPLDWPSDL 118
QY 121 VLAPLPHVNDFTWYKNRKSIGIKPVKNKI--GLSIIPTFNRSRLDITLACLVNQKT 178
Db 119 TLPPLPESTNDYVW-AGRKELDDYPRKQLIIDGLSIVIPYTNRAKILATLACLQNKT 177
QY 179 NYPEFVVVADGSGKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGRLTAKYDFVSI 238
Db 178 IYDEYIVADGSGKENIEEIVREFESLLNKYVRQKDYGYQLCAVRNLGLRAAKYNYVAI 237
QY 239 LDCMAPQOOLWHSYLTLELLEDNDIVLIGRPKYVDTHNITAEQFLNDPYLIESLPETATN 298
Db 238 LDCMAPNPLWQSYMELLAVDDNVALIGRPKYIDTSKHYLDFLSQKSLINEIPEITN 297
QY 299 NNPSITSGNISLDRLEHFKKTDNRLCDSPPRYFVAGNVAFSKWLNKVGWDFEEFNH 358
Db 298 NQVAGKVEQNSVDWRLEHFKNTDNLRLCNTPPRFSSGNVAFKAKWLFRAGWDEEFTH 357
QY 359 WGEDVEFGYRLFAKGFRRVIDGGMALHOEPPGKENETEREAGKSTTLKIVKEKVPYIY 418
Db 358 WGEDNEFGYRLYREGCYFRSVEGAMAYHOEPPGKENETDRAAGKNITVQLLOQKVPYFY 417
QY 419 RKLPIEDSHIHRPLYSIYPAYNCANYIQRVCDSALNOTVVDLEVCICNDGSTDTNLE 478
Db 478 ILQEHYANHPVRPISQKNKGIGASNTAVRLCRGFFYIGQLDSDDFLEPDAVELCLDEFR 537
QY 539 KDKTLACVYTTNRNPNPDGSLIANGYNWPEFSREKLTMTAMIAHFRFTTIRAWHLTDGFN 598
Db 538 KDLSLACVYTTNRNIDREGNLISNGYNWPIYSREKLTSAMICHFRFTTARAWNLTEGFN 597
QY 599 ENTENAVDYDMFLKLSVGVGFKHLNKICYNRVLHGDNSTSIKKLGIOQKNHFVVYNQSLNR 658
Db 598 ESISNVDYDMYLLKLSVGVGFKHINKICYNRVLHGENTSIKKLDIQKENHFKVYNESLSR 657
QY 659 QGINNYNDKFDLDESRYFIENK 682
Db 658 LGIKKYKISPLTNLNECRKYTWEK 681

RESULT 11
US-10-216-289-4
; Sequence 4, Application US/10216289
; GENERAL INFORMATION:

; APPLICANT: NINOMIYA, TOSHIO
; APPLICANT: SUGIURA, NOBUO
; APPLICANT: KIMATA, KOJI
; TITLE OF INVENTION: CHONDROITIN POLYMERASE AND DNA ENCODING THE SAME
; FILE REFERENCE: 226882USO
; CURRENT APPLICATION NUMBER: US/10/216,289
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: JP 2001-244685
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: JP 2001-324127
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: JP 2002-103136
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-216-289-4

Query Match 44.3%; Score 2256.5; DB 6; Length 686;
Best Local Similarity 61.3%; Pred. No. 9.8e-207;
Matches 419; Conservative 112; Mismatches 148; Indels 5; Gaps 3;

QY 1 MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVEFQIIKCKEKLSTNSYVSEDKKNV 60
Db 1 MSTLNQAINLYKNKNYQALSLFEKVAEIVDSVSWVEANIKLCOTALNLSEVDKLNKAV 60
QY 61 CDSLSDIATQLLLSNVKKLTLSSEKSNLKNWKSITGKKSNAEIRKVELVPKDFPKDL 120
Db 61 ID--IDAATKIMCSNAKAIISLNEVEKNEIISKYREITAKKSERAEELKEVEPIPLDWPSDL 118
QY 121 VLAPLPHVNDFTWYKNRKSIGIKPVKNKI--GLSIIPTFNRSRLDITLACLVNQKT 178
Db 119 TLPPLPESTNDYVW-AGRKELDDYPRKQLIIDGLSIVIPYTNRAKILATLACLQNKT 177
QY 179 NYPEFVVVADGSGKENLLTIVQYEQKLDIKYVRQKDYGYQLCAVRNLGRLTAKYDFVSI 238
Db 178 IYDEYIVADGSGKENIEEIVREFESLLNKYVRQKDYGYQLCAVRNLGLRAAKYNYVAI 237
QY 239 LDCMAPQOOLWHSYLTLELLEDNDIVLIGRPKYVDTHNITAEQFLNDPYLIESLPETATN 298
Db 238 LDCMAPNPLWQSYMELLAVDDNVALIGRPKYIDTSKHYLDFLSQKSLINEIPEITN 297
QY 299 NNPSITSGNISLDRLEHFKKTDNRLCDSPPRYFVAGNVAFSKWLNKVGWDFEEFNH 358
Db 298 NQVAGKVEQNSVDWRLEHFKNTDNLRLCNTPPRFSSGNVAFKAKWLFRAGWDEEFTH 357
QY 359 WGEDVEFGYRLFAKGFRRVIDGGMALHOEPPGKENETEREAGKSTTLKIVKEKVPYIY 418
Db 358 WGEDNEFGYRLYREGCYFRSVEGAMAYHOEPPGKENETDRAAGKNITVQLLOQKVPYFY 417
QY 419 RKLPIEDSHIHRPLYSIYPAYNCANYIQRVCDSALNOTVVDLEVCICNDGSTDTNLE 478
Db 478 ILQEHYANHPVRPISQKNKGIGASNTAVRLCRGFFYIGQLDSDDFLEPDAVELCLDEFR 537
QY 539 KDKTLACVYTTNRNPNPDGSLIANGYNWPEFSREKLTMTAMIAHFRFTTIRAWHLTDGFN 598
Db 538 KDLSLACVYTTNRNIDREGNLISNGYNWPIYSREKLTSAMICHFRFTTARAWNLTEGFN 597
QY 599 ENTENAVDYDMFLKLSVGVGFKHLNKICYNRVLHGDNSTSIKKLGIOQKNHFVVYNQSLNR 658
Db 598 ESISNVDYDMYLLKLSVGVGFKHINKICYNRVLHGENTSIKKLDIQKENHFKVYNESLSR 657
QY 659 QGINNYNDKFDLDESRYFIENK 682
Db 658 LGIKKYKISPLTNLNECRKYTWEK 681

```

Qy 755 AFYHKHGVN-----ILLNNDISVYTSNRLLIKTEAHLNSNKNLSQMLNCEYIIFDN 805
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 673 FGQEKTCGNGTAIEMFSIIADND-----TSN-----ASGVIDFVYMGDNTKII----- 716
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 806 HDSLFVKNDSVAYMKKYDVGMNFSALTHDWIEKINAHPPPKKLIKTYFNDNDLRSNNVKG 865
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 717 HDSLLVS-----LLIKREGTTLKSIVQAEW-----EKG 745
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 866 ASQGMFMKYALPHELL--TIKEVITSCQSIDSVPYNTEDIW-----PQFALLIL 914
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 746 -----YTLQENMYTTFDNVFT-----IWAGYTQKYAAFDYNVRL 782
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 915 EKTGHVFNKTSITLYTWPBWRKQLQWNTNEQIOSAKKGKENTPVNKFINSIT 964
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 783 TSREGE-----THFSVVRQNGYIQDTMTSIGNELTPINKIEGNEPT 824
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-10-264-213-152
; Sequence 152, Application US/10264213
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Lubbers, Mark William
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them, and methods for using them.
; FILE REFERENCE: 11000.1043c3
; CURRENT APPLICATION NUMBER: US/10/264,213
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-10-264-213-152

Query Match 5.9%; Score 299; DB 6; Length 331;
Best Local Similarity 27.9%; Pred. No. 2.7e+20;
Matches 87; Conservative 58; Mismatches 131; Indels 36; Gaps

Qy 432 IPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVINKLYGNNPRVR 491
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 MPLVSIIVPIYVNEKYLQRCIDSLFAQTYVNIETLVDDGSADRSLSICKKNAQQDERVR 65
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 492 IMSRPNGGIASAANAYSFAKYYIGOLDSDDYLEDPAVELCKLEFLDKTKLACVYTTNR 551
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 VFSKTNGVDATRNFGVSARGOWISEVPDDYVDPDIETLYGILYKVSNAWMSI-CQHR 1244
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 552 NVNPDGSLIANGYNWPE-----FSREKLTTAMIAHHPMFETIRAWHLTDG-FNE 5999
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 NVYKNGKIQTNLVYEGPAVLDSHTAVKRLLYDDQIDTSVWAKLYPAWFKKIHFPKGRLPE 184
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 600 NIENAVDYDFMLKLESEV---GKPKHLNKKICYNRVLHGDNSTSIKKLGI-----OKKN 6477
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 DI--AATYKTFPLASDSIAGSEAKYSYPPRYNSIVN-DQFSLKHLDLIDMTEQACEVKK 2411
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 648 HFVVVNSLNROGINYY-----NVDKFDLDESRRKYFNKTAEQVEQMDMLDKLKL-IQN 701
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 242 VYDPLQATASORRVLYAYISTLNQNVGDYGOETREQLISKIKELRK--PVLEDHKAPIRD 2999
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 702 KDAKIAVSIFYP 713
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 300 KVAICLIGFSYP 311
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-134-000C-5087
; Sequence 5087, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

```

```
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5087
; LENGTH: 534
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5087

Query Match          5.2%; Score 262.5; DB 5; Length 534;
Best Local Similarity 23.0%; Pred. No. 1.9e-16;
Matches 133; Conservative 100; Mismatches 234; Indels 111; Gaps 22;

QY 260 DNDIVLAGPKRYVD--THNITAEQFLNDPVLIESLPETATNNPSITSKGNISLDWR--- 314
Db 15 NEDIKIFDSIYRDKTNNLTITGALDTIKES--PTFTINNENQYSAYNIQVRLREDV 72

QY 315 -----LEHFKTDNRLCDSPPRYFVAGNVAFSEKWLNK-----VGF 352
Db 73 NQIYQTEPAIEAGFVWTLLEGIKQKKVL-----PFHFOSSAHVITVDPELNKKYVPVPGTE 127

QY 353 DEEFNHGGEDVEGYRLFAGKCFRVIDGMAIHQEPKGENETEREAGKSIILKIVKE 412
Db 128 DKVTRLW-----IKAKKGPKYMAKNGIS--HTQRAKIEKLRNAQ--SYPNWLARN 174

QY 413 KVPYIRKLLPIEDSHIRPLVSIYIPAYNC-ANYIQRVCDSALNQTVDLEVCICNDG 471
Db 175 EVLDI--EAMTQETATPHYQKISIAMPEVYNVEEKWLRCLDSILNQVYTNWELCMADDA 232

QY 472 STD-NLIEVINKLNNDPRVIM-SKPNGGIASNAVAFKAGYIGQLSDDDYLEPDA 529
Db 233 STDPNVKKILTQYQQLDERIVRFREQNGHISEATNSALATATGEFVALLDNDDELAINA 292

QY 530 VELCLKEFLDKTACVYTTNRNVNPDGSLIANGYNWPEFSREKLTITAMIAHFFRMFTIR 589
Db 293 FYEVKVLNENPELDLIYSDDEKIDMDGNRSDPAFK-PWSPDLLLGTNYISHLGVYRS 351

QY 590 AWHLTDFGENIENAVDYMFLKSE---VGKFKHLNLCYNRVLHGDNTSIKK--LGIQ 644
Db 352 ILEETGFRKGYEGSQDYDLVLRFEKTKKERIKHPIKVLXYWRMLPTSTAVDQGSKGYA 411

QY 645 KKHVVVYNQSLNROGINYN-----YDKFDLDESRYKIFPNKTAEOEEMDMKDL 696
Db 412 FEAGLRVQDALVRGNGHATHGAANGLYDVIYDI-ESEK----- 451

QY 697 KLIQNKDAKIAVSIFYPNTLNG---LVKKNLNNIEYNKNIFVILHVDKNHLP----- 747
Db 452 -----LVSIILP-TKNYKDVQRCVSSIIETTYQNYVEIIMADNGSTDPKMHLY 500

QY 748 -DIKKEILAFYHKHQNILLNNDISYYTSNRLIKTEAH 784
Db 501 AEFEQQLPGRFVESIDIPFN---FSTINNRAAKKAH 534
```

RESULT 15

```
US-09-134-000C-6392
; Sequence 6392, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 6392
; LENGTH: 330
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-6392
```

```
Query Match          4.9%; Score 248.5; DB 5; Length 330;
Best Local Similarity 24.9%; Pred. No. 1.8e-15;
Matches 85; Conservative 63; Mismatches 125; Indels 69; Gaps 14;
```

```
QY 432 IPLYSIIPAYNCANYIQRVCDSALNQTVDLEVCICNDGSTDNTLEVINKLYGNPRVR 491
Db 7 MPKISIIIPVYNVEKYLEKCVRSIIAQFTDFELLVDGSPDSSGAMCDQFAEQDQVRK 66

QY 492 IMSPNPGGIASASNAVAFKAGYIGQLSDDDYLEPDAVELCLKEFLK---DKTLACVYT 548
Db 67 VIHRENGGLSDARNAGIATGEVILGFVDSDDYIADDMVELLYTNIVKEDADLSICGIYD 126

QY 549 TNRNVNPDGSLIANGYNWPEFSREKLTITAMIAHFFRMFTIRAWHLTDGFENENIENAVDYD 608
Db 127 VYEGKEPIVKSIIQ-----TFSREALLLILQCN--IISVHA----- 162

QY 609 MFLKLSYGVKFKHLNLCYNRVLHGDNTSIKKIGIQKKNHFWVNOSLNROGIN----- 662
Db 163 -----VNKL-YKRKLFADLRYPK--GYHEDSFIIVDLLSECQKVSIDSTOK 206

QY 663 YNYVDKFDLDESRYKIFENKTAEOEEMDMKDLKLIQNKDAKIAVS---IFYPNTL-- 716
Db 207 YYYHHRMSI--NTEFSDKQFEFTEAWEK-NELKL-KGKGAVIEAAAHQVCFANFLVL 262

QY 717 -----NGLVKK-LNNIIEYNKNIFVILHVDKNHLPDPIKK 751
Db 263 DKILISNAPKKKTKQIVRYLRFENFIFIM---KNKVFTKSRK 301
```

Search completed: January 4, 2003, 02:45:48
Job time : 565 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 4, 2003, 02:18:59 ; Search time 69 seconds
(without alignments)
1344.489 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIAKAYNSNDYELAL.....SAKKGNIPIVKNKIINSITL 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4486.5	88.2	972	T09595	glucuronosyltransf
2	294	5.8	604	E97757	hypothetical prote
3	292.5	5.7	327	AB1211	glycosyltransferas
4	268.5	5.3	696	D95206	hypothetical prote
5	264.5	5.2	323	AD2189	hypothetical prote
6	261.5	5.1	303	A84114	glycosyltransferas
7	258.5	5.1	333	B97168	glycosyltransferas
8	256	5.0	321	AG2188	hypothetical prote
9	254.5	5.0	324	AB2190	hypothetical prote
10	251	4.9	298	B75096	glycosyl transfera
11	241	4.7	333	AH2026	hypothetical prote
12	239	4.7	340	T44330	glycosyl transfera
13	238	4.7	306	T50038	beta-1,3-N-acetylgl
14	238	4.7	316	AE2189	hypothetical prote
15	237.5	4.7	257	E84107	telchuronic acid b
16	237	4.7	344	A70037	capsular polysacch
17	235.5	4.6	318	E71690	minor teichoic aci
18	235	4.6	318	AH2189	hypothetical prote
19	234.5	4.6	336	A97168	glycosyltransferas
20	234	4.6	322	T44647	glycosyl transfera
21	233.5	4.6	1013	AE1876	hypothetical prote
22	232	4.6	333	E97167	glycosyltransferas
23	229.5	4.5	349	D81027	lacto-N-neotetraos
24	229.5	4.5	776	F81289	probable sugar tra
25	229	4.5	389	E81318	probable galactosy
26	228	4.5	392	H69814	hypothetical prote
27	227.5	4.5	333	S70813	glycosyl transfera
28	224.5	4.4	323	H64130	glycosyl transfera
29	223.5	4.4	338	E91190	probable regulator

30	223.5	4.4	338	2	F86037	probable regulator
31	223	4.4	318	2	T50039	beta-1,4-galactosy
32	221	4.3	330	2	AH2188	hypothetical prote
33	221	4.3	334	1	G71153	hypothetical prote
34	221	4.3	338	2	E97083	glycosyltransferas
35	220.5	4.3	445	2	B81289	probable sugar tra
36	220	4.3	329	2	AG0023	probable glycosyl
37	218.5	4.3	318	2	AG2189	hypothetical prote
38	217.5	4.3	344	1	O3ECTH	hypothetical 40.5K
39	217	4.3	367	2	G95948	probable glycosylt
40	216	4.2	390	2	C81318	probable galactosy
41	215.5	4.2	346	2	H81970	lacto-N-neotetraos
42	214.5	4.2	301	2	F95205	glycosyl transfera
43	214.5	4.2	515	2	B81318	probable two-domai
44	214	4.2	250	2	A64099	glycosyl transfera
45	214	4.2	313	2	AI2404	hypothetical prote

ALIGNMENTS

RESULT 1
T09595

glucuronosyltransferase (EC 2.4.1.17) - Pasteurella multocida
N:Alternate names: hyaluronan synthase

C:Species: Pasteurella multocida

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T09595

R:DeAngelis, P.L.; Jing, W.; Drake, R.R.; Achyuthan, A.M.

J. Biol. Chem. 273, 8454-8458, 1998

A:Title: Identification and molecular cloning of a unique hyaluronan synthase from

A:Reference number: Z16757; MUID:98192645; PMID:9525958

A:Accession: T09595

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-972 <DEA>

A:Cross-references: EMBL:AF036004; NID:G3043922; PIDN:AAC38318.1; PID:G3043923

A:Experimental source: strain P-1059; ATCC 15742

C:Genetics:

A:Gene: HAS

C:Function:

A:Description: polymerizes hyaluronan (HA, hyaluronate, hyaluronic acid) polysacchar

C:Keywords: capsule synthesis; glycosyltransferase; hexosyltransferase; hyaluronic a

Query Match 88.2%; Score 4486.5; DB 2; Length 972;
Best Local Similarity 86.6%; Pred. No. 1.2e-238;
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;

QY 1 MNTLSQAIAKAYNSNDYELALFKFEKSAETGRKIVEFQIIKCKEKL----STNS---YVS 53

Db 1 MNTLSQAIAKAYNSNDYQLALKFKFEKSAEYGRKIVEFQITKQEKLSAHPVNSAHLNVN 60

QY 54 EDKNSVCDSSDIATOLLNSNVKLLTSESEKNSLNKWKSTGKKSNAEIRKVELVP 113

Db 61 KEKVNVVCDSPDLATOLLNSNVKLLVSDSEKNTLNKWLKTEKKSENAEVAVALVP 120

QY 114 KDPKDLVLAPLDPHVNDFTWYKNRKSGLGTPVKNKIGLSIIITPTNRSRLDITLACL 173

Db 121 KDPKDLVLAPLDPHVNDFTWYKRRKRLGLKPEHQVGLSVITVTNRPAILSIITLACL 180

QY 174 VNQRTNYPFEVWVADGSKENLLTIVQYEQKLDIKVVRQKDYGYQLCAVRLGLRTAKY 233

Db 181 VNQKTHYPFEVIVTDDGQEDLSPIIQYENKLDIRVVRQKNGFQSAARNMGLRLAKY 240

QY 234 DFVSLDCDMAPQOLVHVSYLELLEDNDIVLIGPKRYVDTHNTATOFNLNDPVLIESLP 293

Db 241 DFIGLLDCDMAPNPLVHVSVAELLEDLTIIGPKRYIDTQHDPKDFLNNSLLESPL 300

QY 294 ETATNNPSTSGNISLDRLEHFKKTDNLRLCDSPFRFVAGNVAFSKENLKNVGF 353

Db 301 EVKTNNSVAAKGEGTSLDRLEHFKKTDNLRLCDSPFRFVAGNVAFSKENLKNVGF 360

QY 354 EEFNHWGGEVDFCYRLFAGKCFRVIDGGMAIHOEPPGKNETEREAGKSITLKIVKEK 413

Db 361 EEFNHGGEDEVGRLLRYSFETIDGIMAYHOEPCKENETDREAGKNTLIDIMEK 420
Qy 414 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCINDGST 473
Db 421 VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCINDGST 480
Qy 474 DNTLEVINKLGNNPRVIMSKPNGGIIASASNAVSPAKGYIIGQLSDDDILEPDAVELC 533
Db 481 DNTLEVINKLGNNPRVIMSKPNGGIIASASNAVSPAKGYIIGQLSDDDILEPDAVELC 540
Qy 534 LKEFLKDKTLCACVYTTNNRNPVDPGSLIANGYNWPEFSREKLTAMIAHFRFTIRAWHL 593
Db 541 LKEFLKDKTLCACVYTTNNRNPVDPGSLIANGYNWPEFSREKLTAMIAHFRFTIRAWHL 600
Qy 594 TDGFNENIENAVDYMFLKSEVGFKHLNKTICYNRVLHGDNTSKKLGIOKKNHFVVVN 653
Db 601 TDGFNEKIENAVDYMFLKSEVGFKHLNKTICYNRVLHGDNTSKKLGIOKKNHFVVVN 660
Qy 654 QSLNRQGINNYNDFDDESKYIFNKTAQYQEMDMKDLKIQNDKAKIAYSIFYP 713
Db 661 QSLNRQGIYINDFDDESKYIFNKTAQYQEIIDLKIKIQNDKAKIAYSIFYP 720
Qy 714 NTLNGLVKKLNNIIEYNKNFIVILHVDKNHLLTPDKKEILAFYKHQHVNILLNNDISY 773
Db 721 NTLNGLVKKLNNIIEYNKNFIVILHVDKNHLLTPDKKEILAFYKHQHVNILLNNDISY 780
Qy 774 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTH 833
Db 781 TSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGMMFSALTH 840
Qy 834 DWIEKINAPPEKLIKTFTFNDNLSRMNVKMGASQGMFKYALPHELLTIKEVITSCQS 893
Db 841 DWIEKINAPPEKLIKTFTFNDNLSRMNVKMGASQGMFTYALAHELLTIKEVITSCQS 900
Qy 894 IDSVPYNTEDIWFOFALLILEKKGHVFNKSTLTLYMPWERKLTQNTNIOQSAKGENI 953
Db 901 IDSVPYNTEDIWFOFALLILEKKGHVFNKSTLTLYMPWERKLTQNTNIOQSAKGENI 960
Qy 954 PVNKFIIINSITL 965
Db 961 PVNKFIIINSITL 972
RESULT 2
E97757
hypothetical protein RC0461 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: E97757
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: E97757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-604 <KUR>
A:Cross-references: GB:AE006914; PIDN:AA02999.1; PID:g15619533; GSPDB:GN00173
C:Genetics:
A:Gene: RC0461
Query Match 5.8%; Score 294; DB 2; Length 604;
Best Local Similarity 24.8%; Pred. No. 1.le-08;
Matches 124; Conservative 71; Mismatches 213; Indels 92; Gaps 19;
Qy 419 RKLPIEDSHIRPLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCINDGSTDN-TL 477
Db 4 RKDKPLNIYH---TLVSIITIPVINGANYKREINSAALQTYKNIIEIIVVNDGSKDNGET 59
Qy 478 EVTNKLGNNPRVIMSKPNGGIIASASNAVSPAKGYIIGQLSDDDILEPDAVE----LC 533
Db 60 ERVALSYGD--KIRYFYKENGCGGCSALNYGKIKMQGYEFSWLSHDDIYYNPKIEHQVDIL 117

Qy 534 LKEFLKDKTLCACVY-----TTNRNVPDGSIIANGYNWPEFSREKLTAMIAHFRMF 586
Db 118 NKLDNKDIIIVGGYELIDKGNLSRYIKPDSVLPINKLN---ISLLPLLRGLIHGCSLLM 174
Qy 587 TIRAWHLTDGFNENIENAVDYMFLKSEVGFKHLNKTICYNRVLHGDNTSKKLGIOKKN 646
Db 175 PAKYFHEVGIFNEALPTTQDYDLWFKIFRVAPIHFDESILIKSRFHSEGS-KKISNNHE 233
Qy 647 N-----HVVVQNQSLNRQGINNYNDFDDESKYIFNKTAQYQEMDMKDLK 697
Db 234 ECVNLWSSFLHETBEEMIKMEGSPYL-----FTRTATFLSNNT-PYKKACDLANTWA 286
Qy 698 LIQNKDAKIAYIFPNTLNGLVKKLNNI-IEYNKNFIVILH-----VDKNHLLTPDIKKE 752
Db 287 KOVLNDTKISVILPYVNRINWAIEAKSVLIQTHKNFEIILDDGSTDDOISETAICKKD 346
Qy 753 -ILAYHKKHQVILLNNDISYTSNRLIKTEAHLNKLNSQLNCEYIIFDNHDSLFV 811
Db 347 KRIKYFHK-----KNEGPAAR-NLGKNAIGKYIAFLDSDDLFY 385
Qy 812 KNDSYAYMKKYDVGMMFSALTHDWIEKINAHPPFKKLIKTYFNDNLSRMNVKMGASQGMF 871
Db 386 K-DKIEIQKPFMEENN-IFSHTSYHKINE-----KGKIESVHSGLF 426
Qy 872 MKYALPHELLTIKEVITSC 891
Db 427 SGNVFP-----QVIQTC 438
RESULT 3
AB1211
glycosyltransferases homolog lmo1090 [imported] - Listeria monocytogenes (strain EGD-
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB1211
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloer
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehl
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:2137279; PMID:11679669
A:Accession: AB1211
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99168.1; PID:g16410492; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1090
Query Match 5.7%; Score 292.5; DB 2; Length 327;
Best Local Similarity 27.5%; Pred. No. 5.9e-09;
Matches 98; Conservative 44; Mismatches 133; Indels 81; Gaps 9;
Qy 433 PLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCINDGSTDNLEVINKLGNNPRVIR 492
Db 3 PLVSIYIPAYNCANYIQRCDVSALNQTVDVLEVCINDGSTDNLEVINKLGNNPRVIR 60
Qy 493 MSPNGGIASASNAVSPAKGYIIGQLSDDDILEPDAVELCFLKDKTLCACVYTTNRN 552
Db 61 FEKNGQATARNFGLDVATGDIYVMVDSDDYISKNLVETCL-DTVQKTNADLVLFYSYN 119
Qy 553 VNPDS-----LIANGYNWPEFSREKLTAMIAHFRFTIRAWHLTDGF--- 597
Db 120 VNQEGKMQYIKRKDGKIVLDAGPTPNKFKYQADLWKG-----SRFPVGYEYEDLGIPLV 173
Qy 598 -----NENIENAVDY---DMFLKSEVGFKHLNKTICYNRVLHGDNTSKKLGIOKKN 647
Db 174 VTLKAKNPVKIQDALYIITDRADSQSNIQOVDFLDVVI---MLENVETELKKLGI----- 227

QY 648 HFVVVNSLNROGINYYNDKFDLDSRKYYIFNKTAETAYEQEEMDMKDL---KLIONKDA 704
Db 228 -----YEEKDQALYLYIEHLYRLVL 249
QY 705 KTAVSIFYPNTLGLVKLNIIIEYKNKFIIVLHVKNHLPDIIKEILAFYHKH 760
Db 250 RKAIIYITNKRKKLIKTIITQIEKFPNMGSPYQAGKLTATLKKKALWLYLHH 305
RESULT 4
D95206
hypothetical protein Sp1771 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: D95206
R:Tetellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95206
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-696 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75845.1; PID:g14973268; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: Sp1771
Query Match 5.3%; Score 268.5; DB 2; Length 696;
Best Local Similarity 20.5%; Pred. No. 3.5e-07;
Matches 101; Conservative 96; Mismatches 181; Indels 115; Gaps 17;
QY 435 VSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSDTNTLEVINKLYGNNPRVRMS 494
Db 7 ITVIVPVYNNVLRKCLDSIIQTQYKNIIEVVVNDGSDASGEICEKSEMDHRIILYIE 66
QY 495 KPNGGIASASNAVSAFAGYIIGDSDDYLEDPAVELCLK---EFLKDKTLACVYTNR 551
Db 67 QENAGLSAARNTGLNNSMGNYVTFVDSDDWIEODYVETLYKTIYEQADIAVGNYSFNE 126
QY 552 N-----VNPDSLIANGYNWPFESREKLTMTAMIAHFRMTIRAWHL-- 593
Db 127 SEGMYFHILGDSYEEKVDNVSIFENLYE---TQEMKSFALISANGKLYKARLFEQLR 182
QY 594 -----TDGFNENIENAVDYMFLKSEVGFKHLNKICY-NRVLHGD----- 634
Db 183 FDIGKLGEDGY-----LNQKVYL-LSE--KVLYLNKSLYAYRIRKGSLSRVWTEKWMH 232
QY 635 -----NTSIKKLGIOKKHNFVVVNSLNROGINYYNDKFDLDSRKYYIFNKTA 695
Db 233 ALVDAMSERITLLANNGYFLEKHLAYIRQMLEVSLAN---GOAGSLSDTATY-----KE 283
QY 686 YOEEMDKLKLIONKDAKIAVSIFYPNTLGLVKLNIIIEYKNKFIIVLHVD-KNH 744
Db 284 FEMKQLLNQLSROESEKKAIVLAANYGVVDVLTIKSICVHNKSIREFYIHSFPNE 343
QY 745 LTPDIKEILAF-----YHKHOVNILLNNDISYTSNRLIKTEAHLNKNKLSQNL 796
Db 344 WIKQLNKRLEKFPDSEIINCRVTSEQIS-CVKSDISYTVFLRYFIADF---VQEDKALYL 398
QY 797 NCEYIIFNDHDSLFVN-DSY-----AYMKKYDVGMN 827
Db 399 DCDLVVTKNLDLFAFDLADYPLAAVYRDFGGRAYFQGEIFNAGVLLVNNFAFWKKENNTQK 458
QY 828 FSALTHDWIEKIN 840
Db 459 LIDVTNWHDKVD 471
RESULT 5

AD2189
hypothetical protein alr3067 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Title: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AD2189
R:KaneKO, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri,
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabat
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2189
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA874766.1; PID:g17132161; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3067
Query Match 5.2%; Score 264.5; DB 2; Length 323;
Best Local Similarity 25.4%; Pred. No. 2e-07;
Matches 82; Conservative 64; Mismatches 116; Indels 61; Gaps 12;
QY 432 IPLVSIYPAYNCANYIQRCDVSALNQTVDLEVCICNDGSDTNTLEVINKLYGNNPRVR 491
Db 10 VPIISVIIPYNGEKTIIETIASVOHOFDIEIIVINDGSDTNTFELVRNIQDN--RLK 67
QY 492 IMSPKNGGIASASNAVSAFAGYIIGDSDDYLEDPAVEL---CLKEFLKDKTLACVYT 548
Db 68 IFSYENGGLPVARNRGITHAVGQFIADIDDLWTTKLELQFAALQEQY-PEAGLAYSMT 126
QY 549 TNRNVNPDGS-----LIANGYNWPFESREKLTMTAMIAHFRMTI 588
Db 127 YKFAEADSYADESNFSAGDYVAELLIKNFQNGSN-PLIRAAIDSVGL----- 176
QY 589 RAWHLTDGFNENIENAVDYMFLKSEVGFKHLNK--ICYNRVLHGDNTSIKKLGIOKK 646
Db 177 -----FDPTLKSCEDWDFYLRLAQWQFALYKKAQIYRQ---SPTAMTSKLDYMEK 225
QY 647 NHFVVVNSLNROGINYYNDKFDLDSRKYYIFNKTA--YQEEMDKLKLIONKDA 704
Db 226 YSSIVIERAFNAAPQQLHLKK-----QSLAWYVKFTAQCLKYNSHKLADIKLAA-KRL 279
QY 705 KIAVSIFYPNTL---NGLVKKL 723
Db 280 KMAITLYPKNLLLEDYTHGLIRKL 302
RESULT 6
A84114
glycosyltransferase BH3713 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: A84114
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masul, N.; Fujii, F.;
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A84114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-303 <STO>
A:Cross-references: GB:AP001519; GB:BA000004; MID:g10176109; PIDN:BA807432.1; GSPDB
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3713
Query Match 5.1%; Score 261.5; DB 2; Length 303;
Best Local Similarity 26.5%; Pred. No. 2.7e-07;
Matches 90; Conservative 60; Mismatches 95; Indels 95; Gaps 15;
QY 435 VSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSDTNTLEVINKLYGNNPRVRMS 494

Db 3 VSIIPHNRAKLLKRALESLNQTNYKNIIEVIVVSDGSTNT-DIVMDYKRDNRVNFIS 61
Qy 495 -KPNGGIASANAASFAGYIIGQLSDDDYLEDPAVELCLKEFLKDKTLACVYT----- 548
Db 62 YHPAKGGYARNTGKNAKGFIAFLDSDDDMPDKLELQIKFNNQANVGLVYTGVEII 121
Qy 549 ---TNRNVNPDGSLIANGYNWPE---FSREKL-----TTAMTAHHPFMTTIRAWHLTD 595
Db 122 YNFNKRNI-----KYSLPKTKGNLSKEILVANCIGTTSV-----MVRKNLITE 166
Qy 596 --GFENIENAVDVMFLKSEVGKFKHLKICNYNRVLHGDNSTIKKLGIOKKNHFFVYVN 653
Db 167 CGMFEDEKLARQDYDLWI-----RVC-----QKTLVGVVN 196
Qy 654 QSLNRQGINNYNDKFDL-DESKY-----IFNKTAE-----YQEMDMKL 694
Db 197 KPL-----VRYNYNTTKQISDDIKYSAIEYIDNKYVDLYSKVSEETRRKRRHSMTMLI 252
Qy 695 DLKLTIONKDAKIA-----VSIFYPNTLNGVLKLNIIIEY 729
Db 253 VNKALRNSPKVARAYLKNSEFLKRLTLTALIMYLSFLKY 292

RESULT 7
B97168
glycosyltransferase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97168
R: Nolling, J.; Brennet, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J: Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97168
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK80133.1; PID:g15025169; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2175

Query Match 5.1%; Score 258.5; DB 2; Length 333;
Best Local Similarity 26.4%; Pred. No. 4.4e-07;
Matches 81; Conservative 54; Mismatches 125; Indels 47; Gaps 11;

Qy 434 LVSIIYPAYNCANYIQRCDVSALNQTVVDLEVCICNDGSDTNTLEVINKLYGNPNRVRIM 493
Db 1 MISVIMPVYNCBKYLESTIESILKQTYRDFEIIVNDGSDNKSIDIINKYANDDNKRIVV 60
Qy 494 SKPNG-GIASASNAASFAGYIIGQLSDDDYLEDPAVELCLKEFLKDK---TLACVYTT 549
Db 61 SRDNNMGVYSLNEGIDRAGSVARMADDIALPERFERQIEYLNKNDVDILACKVEA 120
Qy 550 NRVNPDGSL-IANGYNWPEFSREK-----LATAMTAHHPFMTTIRAWHLTDGFENIEN 603
Db 121 FGDVSRQKLERHWHYVNDLNNSESTESLFLENCYIAHPVSMKMSVLKALGGYNLYKR 180
Qy 604 AVDYDMFLKSEVG-KFKHLKICNYNRVLHGDNSTIKKLGIOKKNHFFVYVNSLNROGIN 662
Db 181 TEDYNLWLRAIAGKYIAMLEEKMKIRLHNS-----KIHDA-- 219
Qy 663 YYNDRFDDLDDESKYIFNKTAEQEOMDMLDKLI---QNKDAKIA---VSIFYPNT- 715
Db 220 -EGFSSIRDIQSR-----LEYVKEKLKDKDFSVINGASNGGKIAYEKIKEVFPNAK 271
Qy 716 LNLVVRK 722
Db 272 LNGYIDK 278

RESULT 8
AG2188
hypothetical protein alr3062 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AG2188
R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2188
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA874761.1; PID:g17132156; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3062

Query Match 5.0%; Score 256; DB 2; Length 321;
Best Local Similarity 27.5%; Pred. No. 5.8e-07;
Matches 78; Conservative 50; Mismatches 118; Indels 38; Gaps 10;

Qy 432 IPLVSIYPAYNCANYIQRCDVSALNQTVVDLEVCICNDGSDTNTLEVINKLYGNPNRVR 491
Db 1 MPKVSWVIYPAYNAMPYLPETLESVLRQTYVHDFEVVVVNDGSSDNTTEEWVSQIL--DPRLK 58
Qy 492 IMSKPNGGIASASNAASFAGYIIGQLSDDDYLEDPAVELCLKEFLKDKTLACVYTTNR 551
Db 59 LISQANQGLAGARNTGIVNASGEYIAFDADDIWEPTKIAKQSVLDENPTVGLVYTWVA 118
Qy 552 NVNPDGSLI-----ANGYNWPEFSREKLTAMTAHHPFMTTIRAWHLTDGFENIEN 604
Db 119 YDEQKSGTKITFKNQVEGYVWPQLTEHNVCEGVALVRRVCFEKMGL---FDRNLGYS 175
Qy 605 V-DYDMFLKSEVGKFKHLNK-ICYNRVLHGDNSTIK-----KLGIQKKNHFFVYVNS 655
Db 176 VEDWDMWLRIATSYDFKVVKEALVYR--ORSNSASKNWEAMAHSAFIVIEKAFATASQD 233
Qy 656 L-----NROGIN-----YNYDKFDDLDDESKYIFNKTAIEY 686
Db 234 LQVLNKSXGFTYLCCLAWKPLQSFQK--DYQKSREFCQQAIVY 275

RESULT 9
AB2190
hypothetical protein alr3073 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AB2190
R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2190
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA874772.1; PID:g17132167; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3073

Query Match 5.0%; Score 254.5; DB 2; Length 324;
Best Local Similarity 26.2%; Pred. No. 7.1e-07;
Matches 101; Conservative 63; Mismatches 127; Indels 95; Gaps 16;

Qy 432 IPLVSIYPAYNCANYIQRCDVSALNQTVVDLEVCICNDGSDTNTLEVINKLYGNPNRVR 491
Db 1 MPKVSWVIYPAYNAMPYLPETLESVLRQTYVHDFEVVVVNDGSSDNTTEEWVSQIL--DPRLK 58

Db 1 MPKISVIIIPAYNAERTILETINSVNLQTFSDLEIIIVINDSGDRTVEVLQNV--DDARLK 58

QY 492 IMSKPNGGIASNAVSPAKYIYGQLSDDDYLEPDAVELCIGKFLKDKTKLACVYTTNR 551

Db 59 VYSYENSASGARNHGICSHAVGDFISFDADDLWTPDKLEQLSALNNHPEAGVAYSWTY 118

QY 552 NYNPDGSLIANGYNNPFSRREKLTMTAMIAHFRMTIRAWHLTDGFNENIENAVDYDMEL 611

Db 119 TIDDKGELLKPE--PLYEGNYVTDLLAN-----FLTNGSNPLIRKAA----- 160

QY 612 KLSVEGKFKHLNKCYNRVLHGDNTSIIKK-----LGTKKNHFVVVNQSLNRQGINYY 664

Db 161 -IASICEF-----DITLRSGEDWYWLRLAYKWPVVVRQ-----HQILY- 199

QY 665 NYDKFDLDESRYIFNKTAEOEEMDKLKL-IQNKDAIYVIFYPNTLGLVK-K 722

Db 200 -----RRSVTSKSFQIQ-----IIREASLAIDKAMKV-----LPLEQYLKHS 239

QY 723 LNNIIEYNNKIFVIIHVDKNHLPDIIK-----KEILAFYHKHVNIL 765

Db 240 LSNIRYNNVELYDSIN---NNSVDIKYVIGNLLSYIRSRQTLKEI--YYKLIIL 294

QY 766 LNDISYTSNRL----IKTEAHLN 788

Db 295 LVIVLSPKLSRLLOFIKSKOMKML 320

RESULT 10

B75096

glycosyl transferase PAB0772 - *Pyrococcus abyssi* (strain Orsay)

C:Species: *Pyrococcus abyssi*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: B75096

R:anonymous, Genoscope

A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure submitted to the EMBL Data Library, July 1999

A:Reference number: A75001

A:Accession: B75096

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <K>

A:Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB50071.1; PID:g545858

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0772

C:Superfamily: *Neisseria meningitidis* glycosyl transferase A

Query Match 4.9%; Score 251; DB 2; Length 298;

Best Local Similarity 24.3%; Pred. No. 9.8e-07;

Matches 81; Conservative 66; Mismatches 136; Indels 50; Gaps 8;

QY 433 PLVSIYIPAYNCANYIQRVDSALNOTVVVDLEVCINDGSTNTLEVINKLYGNPNRVR- 491

Db 4 PIVSVIIPYNNRANLLRAIAISVLRKFDLIVVDDASTONTPEVSEI--EDGRIRY 61

QY 492 IMSKPNGGIASNAVSPAKYIYGQLSDDDYLEPDAVELCIGKFLKDKTKLACVYTTN 550

Db 62 IRLKNSGGPIARNIGIKKAKGRFIALDDDEWLPHRLEVQVRKFNELGKEGVYGGF 121

QY 551 RVNPDGSLIANGYNNPFSRREKLTMTAMIAHF-----RMFTIRAWHLTDGFNENIENAV 605

Db 122 YVVSQDGRIL--GKRLPK-HRGDIYSHLLKENFIGSPTLLIRRECFKAGLFDPRLESSQ 178

QY 606 DYDMFLKLEVEGKFKHLNKCYNRVLHGDNTSIIKKLGIOKKNHFVVVNQSLNRQGINYYN 665

Db 179 DWDMMRLIARYKFDYVDIEIAKYVYHVGQISF----- 211

QY 666 YDKFDLDESRYIFNKTAEOEEMDKLKLQNKDAIYVIFYPNTLGLVKLN- 724

Db 212 -----NMKKYIPGRBLIRKHLDIWKNPKILSHLSQMLLLLSNNTGKGLKLYY 263

QY 725 --NIEYNNKIFVIIHVDKNHLPDIIKKEILA 755

Db 264 STAIAPLNLENYMIILLKALDSRTVEYIKRILS 296

RESULT 11

AH2026

hypothetical protein all1766 [imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp.

A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AH2026

R:Kanakok, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irlig Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabat DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Nostoc* sp. strain PCC 7120

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2026

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073465.1; PID:g17130856; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1766

Query Match 4.7%; Score 241; DB 2; Length 333;

Best Local Similarity 26.4%; Pred. No. 4.1e-06;

Matches 70; Conservative 53; Mismatches 94; Indels 48; Gaps 7;

QY 432 IPIVSIYIPAYNCANYIQRVDSALNOTVVVDLEVCINDGSTNTLEVINKLYGNPNRVR 491

Db 8 LPIISVIIPAYNCEKTIKTIDSVLDQSFDFELIVINDGSQDATLIVSQI--EDSRIK 65

QY 492 IMSKPNGGIASNAVSPAKYIYGQLSDDDYLEPDAVELCIGKFLKDKTKLACVYTTNR 551

Db 66 IFSFENAGGVNVRNGLNLAAGKFIPLDADDIWTNPKLESQALNHPGFHVAYSWTD 125

QY 552 NVNPDGSLIANG-----YNNPFSRREKLTMTAMIAHFRMTIRAWHLT 594

Db 126 YIDEDGNFLISGRVTLNGDVYKFLFNFLNGSNPLICKEALIA-----L 172

QY 595 DGFNENIENAVDYDMFLKLEVEGKFKHLNKCYNRVLH--GDNTSIIKKLGIOKKNHFVV 652

Db 173 GGFDESLKAAQDWMRLANKYSFV---AVPYQIILRVSSNSLSNLVRQEKACKQVL 229

QY 653 -----NOSLNRRQGINYYN 666

Db 230 EKAYQARPAIGNHILHLSIANLYKY 254

RESULT 12

T44330

glycosyl transferase homolog [imported] - *Vibrio cholerae*

C:Species: *Vibrio cholerae*

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T44330

R:Yamasaki, S.; Shmizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Gene 237, 321-332, 1999

A:Title: The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are

A:Reference number: Z22749; MUID:99453293; PMID:10521656

A:Accession: T44330

A:Status: preliminary;

A:Molecule type: DNA

A:Residues: 1-340 <YAM>

A:Cross-references: EMBL:AB012957; NID:g4115688; PIDN:BAA33634.1; PID:g3721684

A:Experimental source: strain O22

C:Genetics:

A:Note: wblC

C:Superfamily: *Neisseria meningitidis* glycosyl transferase A

Query Match 4.7%; Score 239; DB 2; Length 340;

Best Local Similarity 25.4%; Pred. No. 5.4e-06;

Matches 78; Conservative 56; Mismatches 113; Indels 60; Gaps 11;

QY	433	PLVSIYPAYNCANIQCRCVDSALNQTVVLEVCICNDGSDTNTLEVINLKLYGNPNRVRI	492
Db	6	PTISVIMSVNGEKYLAQALIESIILNQTSDFEFTIIVDGGSDTSLSIQATYMDKDDRLVL	65
QY	493	MSKPNGGSIASNAVAFAKYYIGQLDSDDDYLEPDPAVELCLKEFLKDKTLACVYTTNRN	552
Db	66	ISRVNKGLPYSLNEAISVSKANYIARMDADDISLPERLE-----TQLVWENNPD	115
QY	553	VNPDGSLIANGYNWPEF--SRKLTMTAMIAHF-----RMFTTIAWH-----LTD	595
Db	116	IGVCGTL---AYLFRETSPKNNKMCHPEDHDSLIIRLFSVCFIHPVVMIRKSVLDQLDY	172
QY	596	GFNENIENAVDYDFELKLSVGKFKHLNKK--ICYNRVLHGDNSTSKKLGIOKKKHFEVVN	653
Db	173	VYNENFRNSQDYELWLSRIAETRTFTYIOKPLLFRTDPDGITSKVNHDGFNKR--FPLVS	230
QY	654	Q-----SLN-----RQGIN---YYNDFKDDLDESRRKYIFNKTAEYQEEMD	691
Db	231	QVQTKOLEKLGKLNDEEMIMHFRGLNAEMLFLNAKADDV---KYVFNSILNANKKKTQ	286
QY	692	MLKDKL 698	
Db	287	IFHQKKL 293	
RESULT 13			
beta-1,3-N-acetylglucosaminyltransferase [validated] - Streptococcus pneumoniae			
C:Species: Streptococcus pneumoniae			
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 31-Dec-2000			
C:Accession: T50038			
Mol: Microbiol. 26, 197-208, 1997			
A:Title: Capsular polysaccharide synthesis in Streptococcus pneumoniae serotype 14: mole			
the biosynthesis of the tetrasaccharide subunit.			
A:Reference number: #25028; MUID:98043549; PMID:9383201			
C:Accession: T50038			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-306 <KOL>			
A:Cross-references: EMBL:X85787; PIDN:CAAS9781.1			
A:Experimental source: isolate NCTC 11902; serotype 14			
C:Genetics:			
A:Gene: cps14I			
C:Function:			
A:Description: necessary for the addition of the third sugar in the synthesis of the typ			
A:Pathway: capsular polysaccharide biosynthesis			
Query Match 4.7%; Score 238; DB 2; Length 306;			
Best Local Similarity 25.5%; Pred. No. 5.3e-06;			
Matches 84; Conservative 52; Mismatches 97; Indels 96; Gaps 14;			
QY	435	VSIYIYPAYNCANIQCRCVDSALNQTVVLEVCICNDGSDTNTLEVINLKLYGNPNRVIMS	494
Db	4	VSIYIYPAYNAETIKNCVDSALKQNLSEVLIVNDGSDNSTSKILEQ-YGDNQPMWIFH	62
QY	495	KPNGGSIASNAVAFAKYYIGQLDSDDDYLEPDPAVELCLKEFLKDK---TLACVY----	547
Db	63	QVNWGVSARNVGLSVASGEYVFFLSDSDILDEGMLSK-MYQFAKSNKIDLLSCWHKEPS	121
QY	548	TT--NRNVNPDGSLIANGYNWPEFSREKLTUTAMIAHF-----RMFTTIAWHLTD	595
Db	122	TTQYGGNDSSASFAR-----TKEEIGNHFVDIFPRSAKALFLRRRIE--- 166	
QY	596	GFNENIENAVDYDMFKLSEVGKFKHLNKCICYNRVLHGDNSTSKKLGIOKKKHFEVVNOS	655
Db	167	-----ENNIAFSTMSLGEDMSF-----VC-----QYLMVSR	194
QY	656	LNRRQGINYYNDFKDDLDESRRKYIFNKTAEYQEEMDKLKLQNK--DAKTAVSIFYP	713
Db	195	IAVIDGLYTYIQNVNQSLSKRVSN-----IENSLMQNLMDQLLEV---YP	240
QY	714	NTNLGLVKK-----LNNITEYN	730
Db 241 KIEENYIKQHMDFRYLASLYVNNLFKFD 269			
RESULT 14			
AE2189			
hypothetical protein alr3068 [imported] - Nostoc sp. (strain PCC 7120)			
C:Species: Nostoc sp.			
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120			
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002			
C:Accession: AE2189			
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriga			
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata			
DNA Res. 8, 205-213, 2001			
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium			
A:Reference number: AB1807; MUID:21595285; PMID:11759840			
A:Accession: AE2189			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-316 <KUR>			
A:Cross-references: GB:BA000019; PIDN:BAB74767.1; PID:gl7132162; GSPDB:GN00179			
A:Experimental source: strain PCC 7120			
C:Genetics:			
A:Gene: alr3068			
Query Match 4.7%; Score 238; DB 2; Length 316;			
Best Local Similarity 24.5%; Pred. No. 5.5e-06;			
Matches 78; Conservative 50; Mismatches 109; Indels 82; Gaps 9;			
QY	432	PLVSIYIYPAYNCANIQCRCVDSALNQTVVLEVCICNDGSDTNTLEVINLKLYGNPNRVYR	491
Db	1	MPTISVIIIPAYNAETILETISVQQQTFSDFDLIINDGSTDRTLEIQNI--RDERLK	58
QY	492	IMSKPNGGSIASNAVAFAKYYIGQLDSDDDYLEPDPAVELCLKEFLKDKTLACVYTTNR	551
Db	59	IFSYENGGLCTARNRIGISHASGEFTAFDLADDLWTHDKLEQL-----TAL	104
QY	552	NVNPDSGLIANGYNWPEFSREK-----LVTAMIAHF-----RMFTIRA	590
Db	105	QQHPEAGV---AYSWTYFMEDEQCKSSIPGVSLFFEGDVQAHLLVNNFLASGNSNPLIRKQA	161
QY	591	WHLTDGFENIENAVDYDMFKLSEVGKFKHLNKCICYNRVLHGDNSTSKKLGIOKKHNFV	650
Db	162	IESVGEFDSNCMGCAWDYWLRSATW-----NFV	191
QY	651	VV--NOSLRQGINYYNDFKDDLDESRRKYIFNKTAE-YQEEMDKLKLQNKDAKIA	707
Db	192	VVRKHQIFRYQSATSMSSTKVKNMEDDGLFVVEKTFQSVKPELQYLNQSL-----	242
QY	708	VSIFYPNTNLGLVKKLNNI 726	
Db	243	AWIYQYSTQOYLKQNVNNI 261	
RESULT 15			
E84107			
teichuronic acid biosynthesis tuag [imported] - Bacillus halodurans (strain C-125)			
C:Species: Bacillus halodurans			
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001			
C:Accession: E84107			
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;			
Nucleic Acids Res. 28, 4317-4331, 2000			
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans			
A:Reference number: A83650; MUID:20512582; PMID:11058132			
A:Accession: E84107			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-257 <STO>			
A:Cross-references: GB:AP001519; GB:BA000004; MID:g10176109; PIDN:BAB07380.1; GSPDB:			
A:Experimental source: strain C-125			
C:Genetics:			
A:Gene: tuag			

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 00:40:24 ; Search time 66 Seconds
(without alignments)
606.434 Million cell updates/sec

Title: US-09-842-484A-2

Perfect score: 5089

Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SAKKGENIPVKNFIINSITL 965

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	245.5	4.8	909	1 Y4GI_RHISN	P53465 rhizobium s
2	224.5	4.4	323	1 YF78_HAEIN	Q57287 haemophilus
3	217.5	4.3	344	1 Y1BD_ECOLI	P11290 escherichia
4	214	4.2	250	1 Y868_HAEIN	Q57022 haemophilus
5	212.5	4.2	900	1 G6AB_BACSD	P46918 bacillus su
6	199.5	3.9	322	1 Y866_ANASP	P22639 anabaena sp
7	197.5	3.9	1275	1 RFBC_MYXXA	Q50864 myxococcus
8	193.5	3.8	299	1 Y025_MYCPN	P75086 mycoplasma
9	191.5	3.8	446	1 GGAA_BACSU	P46917 bacillus su
10	189	3.7	1174	1 EX5B_BUCAL	P57529 buchnera ap
11	183	3.6	1169	1 EX5B_BORBU	O51578 borrelia bu
12	177.5	3.5	970	1 Y087_BUCAL	P57189 buchnera ap
13	176.5	3.5	298	1 Y025_MYCGE	P47271 mycoplasma
14	176	3.5	1442	1 DP03_UREPA	Q9pqb4 ureaplasma
15	175.5	3.4	1162	1 BXEN_CLOBU	Q06366 clostridium
16	174.5	3.4	2710	1 TOXA_CLODI	P16154 clostridium
17	174	3.4	2339	1 RPI1_PLAFA	P27625 plasmodium
18	173.5	3.4	986	1 EPIB_STAEP	P30195 staphylococ
19	170.5	3.4	279	1 WCNA_ECOLI	P77414 escherichia
20	168.5	3.3	333	1 REBV_SALTY	P26401 salmonella
21	168	3.3	348	1 EXOO_RHIME	P33697 rhizobium m
22	168	3.3	1162	1 BXEN_CLOBO	P46082 clostridium
23	166.5	3.3	1738	1 YCF1_EPIVI	Q00383 epifagus vi
24	165.5	3.3	1111	1 KIP1_YEAST	P28742 saccharomyc
25	165	3.2	297	1 Y060_MYCGE	P47306 mycoplasma
26	163	3.2	2077	1 TEGU_HSV6G	P30002 human herpe
27	163	3.2	2077	1 TEGU_HSV6U	P52340 human herpe
28	160.5	3.2	1005	1 RA50_METJA	Q58718 methanococc
29	159	3.1	900	1 SYA_MYCGE	P47534 mycoplasma
30	157	3.1	2748	1 NMU1_YEAST	Q00402 saccharomyc
31	155.5	3.1	1202	1 RMP2_YEAST	Q02773 saccharomyc
32	155	3.0	1684	1 DPOL_METJA	Q58295 methanococc
33	154	3.0	256	1 SPSA_BACSU	P39621 bacillus su

34	154	3.0	290	1 YA57_METJA	Q58457 methanococc
35	154	3.0	2136	1 YCF2_MARPO	P09975 marchantia
36	153.5	3.0	2869	1 RBP1_PLAVB	Q00798 plasmodium
37	153	3.0	299	1 Y060_MYCPN	P75042 mycoplasma
38	152.5	3.0	648	1 NTP1_AMEPV	P29814 amsacta moo
39	152.5	3.0	667	1 Y366_MYCGE	P47606 mycoplasma
40	152	3.0	851	1 MCEL_FOWPV	Q01584 fowlpox vir
41	150	2.9	1197	1 DPMO_PODAN	Q01529 podospora a
42	149.5	2.9	294	1 YG96_HAEIN	Q48214 haemophilus
43	149.5	2.9	1324	1 CUT3_SCHPO	P41004 schizosacch
44	148.5	2.9	1750	1 Y832_METJA	Q58442 methanococc
45	148.5	2.9	1928	1 MYS1_YEAST	P08964 saccharomyc

ALIGNMENTS

```
RESULT 1
Y4GI_RHISN
ID Y4GI_RHISN STANDARD; PRT; 909 AA.
AC P53465;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical 102.8 kDa protein Y4GI.
GN Y4GI.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Broughton W.J., Rosenthal A.,
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- SIMILARITY: STRONG TO M_XANTHUS O-ANTIGEN BIOSYNTHESIS PROTEIN
CC RFBC (NOT TO BE CONFUSED WITH ENTEROBACTERIAL RFBC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000074; AAB91683.1;
DR HSP; P39621; LOGO.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 2.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 909 AA; 102827 MW; 3985D69722F43962 CRC64;
```

```
Query Match 4.8%; Score 245.5; DB 1; Length 909;
Best Local Similarity 20.8%; Pred. No. 4.2e-06;
Matches 159; Conservative 129; Mismatches 257; Indels 227; Gaps 38;

QY 20 LKFEKSAETGYGRKIVFQIKCKEKLSTNSYVSVSEDKNSVCDSLLDIATQLLSNVKKL 79
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 VEFPALAGESYVRDATQGVK---RLELAETIEERKAET-----VLVQOKLQEARRR 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 80 TLSESEK-----NSLKNKWKSIITKKSNAIRKVELYVPKDFPPKDLVLAPLDHVNDET-- 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 PLQLKRLKLVFNMLRAAAKASPLPSRTAERFRSAARD-----PMRDDLTQLSGQ 332
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 -----WYKNRKKSIG-----IKPVKNKIGLSIIITPTNRSRILDTLACLNVQK 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 333 GFMTYEAVVRGWGKQKQALAGRLSELVRRRLQNGPLTISVVVFPVNPDPALLVEMIESVRAQ 392
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 TNPFFVWVADDSKE-NLLTIVQKY-EQKLDIKYVRQKDYQLCAVRNGLRLTAKYDF 235
```


DR	ENBL; AF031959; AAC32401.1; ALT_INIT.
DR	PfR; B35391; B35391.
DR	InterPro; IPR001173; Glycos_transf_2.
DR	Pfam; PF00535; Glycos_transf_2; 1.
DR	KW Hypothetical protein; Transferase; Glycosyltransferase;
KW	Complete proteome.
SQ	SEQUENCE 322 AA; 36388 MW; 753C2FB593327D968 CRC64;

Query Match	3.9%; Score 199.5; DB 1; Length 322;
Best Local Similarity	26.4%; Pred. No. 0.00034;
Matches 81; Conservative 48; Mismatches 101; Indels 77; Gaps 16;	

QY	435 VSIYIPAYNCANYIQRCDVSALNQTVVDLEVCICNDGSDTNTLTVEINKLYGNPN-RVIRM 493
Db	: :
Db	3 ISVIISINRYARLSRAINSIAQTSHSDIBIVIVDDGSTNSRDVITQLQEAPDKIKPI 62
QY	494 SKPNGGLASNAAVSFAKGYIGOLDSDDYLDLPDAVELCLKEFL----- 538
Db	: :
Db	63 FQANQGOGGFAGNAGFAATGEVAFDLADDWKPKLQRIIVEVFQTSDDVGVMHLLDIID 122
QY	539 -KDKTLACVYTTRNNVPD--GSLI---ANGYNWP-----EFSREKLTTAMIAHHFRMFTI 588
Db	
Db	123 GNDKTIDAOSTQGPLSEADSLVILQTGNACFPPTSGLAYRRVLE-----KVPEI 174
QY	589 -----RAHWHTDGTGENENIENVADY-DNFLKLSEVGKFKHLNK-ICYNRVLHGDNWTISI--- 638
Db	
Db	175 DPVKRIW--ADG-----CIITYCTAFL-----GKIQLQENLAYYRI-HGANNHMSAA 219
QY	639 -----KKLGIOKKNHFFVWNOSLNROGINNYNDKFDDLDESRYIFNKTAEYQEE 689
Db	
Db	220 SATSEQSAKSQAIGEMTNQY--INDFLVRIGY-----ARVDSLRLQYRRTKYYQRS 270
QY	690 MDMLKDL 696
Db	
Db	271 QWDLRV 277

RESULT 7	
RFCB_MYXXA	
ID	RFCB_MYXXA STANDARD; PRT; 1275 AA.
AC	Q50864;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	01-NOV-1997 (Rel. 35, Last annotation update)
DE	O-antigen biosynthesis protein rfbC.
GN	RFCB.
OS	Myxococcus xanthus.
OC	Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
CC	Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX	NCB_I_TaxID=34;
[1]	SEQUENCE FROM N.A.
RP	STRAIN=DK6640.
RC	MEDLINE=96198166; PubMed=8626291;
RX	Guo D., Bowden M.G., Pershad R., Kaplan H.B.;
RT	"The Myxococcus xanthus rfbABC operon encodes an ATP-binding cassette
RT	transporter homolog required for O-antigen biosynthesis and
RT	multicellular development";
RL	J. Bacteriol. 178:1631-1639(1996).
CC	-!- FUNCTION: INVOLVED IN O-ANTIGEN BIOSYNTHESIS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
ENBL; U36795; AAB05019.1; --	
DR	InterPro; IPR001296; Glycos_transf_1.
DR	InterPro; IPR001173; Glycos_transf_2.
DR	Pfam; PF00534; Glycos_transf_1; 1.

DR Pfam; PF00535; Glycos_transf_2; 2.
KW Lipopolysaccharide biosynthesis.
SQ SEQUENCE 1275 AA; 139596 MW; 3AF9662A10A140F1 CRC64;

Query Match 3.98; Score 197.5; DB 1; Length 1275;
Best Local Similarity 20.98; Pred. No. 0.0026;
Matches 84; Conservative 71; Mismatches 176; Indels 71; Gaps 14;

Qy 427 SHIHRPLSIYIPAYCAN-YIORCVDSALNQTWVLEVCICNDGST-DNTLEVINKLY 484
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 313 ARLSRRPLISLVTVPVDASEAFLECLASVSQVYADWEWLVLDDASTAPHLARILEAA 372
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 485 GNPNRVRIM-SKPNGGTASNAAVSAFKYYIQGLSDDDYLEPDANVELCKELDKTKTL 543
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 373 ERESRIIRLVLTASSBGDTARATNEGFACRCDFGVGLGAEDTSLSPHALAEVALAFAOPEL 432
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 544 ACVVTTNRNVNPDCSLLANGYNRPFEFSREKLTTAMIAHHFRFMETIRAWHLTDGNGENIE 603
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 433 ALLTYDEGLDAQHRSAPFEK-PDWSPLLURSYDYVRHFLLVVRRETIAQVGGLREGFDG 491
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 604 AVDYDMFLKISE-VGKPKHLNKICYNRVYLHGDNYSIKK--LGIOKKNHFFVVMQSINRQ 659
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 492 AQGHDLMLRLSEANSIGHITEPLYH-AREGSASAARGAGLDYATKAGVRALSEHLARQ 550
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 660 GINYNYNDKDDLDESRRKYIFNKTAEYOEMDMCLKLIKQNKDAKIIVSYPTNLNGL 719
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 551 G-----ESAEV-----TSAPIQRYVRYP----- 569
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 720 VKLLNNIIYEKNKFVILHVKKHLPDIKKEILA-FYHKHQVNILLNND----- 769
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 570 -----VRGPVKVIIVPFKRDPLDRLTVDSLQAOTRYPHFVEVLVSNNSNTRPETFAL 622
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 770 ISYTSNRLIK-TEAHLSN---INKLSQLNCEYVIIFDNHD 807
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 623 LEQWVDPRLVLYKTWDHPFNYPAINNWAAKQASGELLLELNND 664
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 8
Y025_MYCPN STANDARD; PRT; 299 AA.
ID Y025_MYCPN
AC P75086;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative glycosyl transferase MG025 homolog (EC 2.-.-.-)
DE (B01_orf299v).
GN MFN028 OR MF126.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
ON NCBI_TaxID=2104;
RX [1]
PP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-37105885; PubMed=8948633;
RA Hammerleisch R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae".
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC STRONG, TO M.GENITALIUM MG025.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000015; AAB95774.1; .
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.

KW Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 299 AA; 34935 MW; 9A3AF2F099283565 CRC64;
Query Match 3.8%; Score 193.5; DB 1; Length 299;
Best Local Similarity 24.5%; Pred. No. 0.00066;
Matches 74; Conservative 38; Mismatches 127; Indels 63; Gaps 10;
QY 434 LVSIYIPAYNCANYIQCVDLSAL--NOTVVDLEVCINDGSDTNTLEVINKLYGNPRVR 491
DB 6 LFTVLIPTNGQVYIPRALDLSLLQGEYFTKTQVLIVDSDGTDNTKTPVEYPTQOYSN 65
QY 492 IMSFPGNGIASASNAV--SFAKYYIGOLDSDDDYLEDPAVELCLKEFLKDKTLACVVT 549
DB 66 YLEKPNMGWGVNPFVQKONLAKQGYITVLDSDDYFLANAFQVAAHFGHDMIVSAFYC- 124
QY 550 NRVNPDGSLIANGYNWPEFSREKLTTAMIAH---FRMFTIRAWHLTDGFNENIENAVD 606
DB 125 --YISPKRRRELKPYEGKTGVIEQTKLRTPHSQPLAKFYRHEIFHLDLPLKEKL---FY 179
QY 607 YDMFKLSEVGKFKHLNKC-----YNRVLHGDNTSIKGLGKIQKHHFVVYNQSLNRQGI 661
DB 180 QDLCLHNAINKVQSVFVCEPLAVYATRFQCNSTMP----- 217
QY 662 NYNYTKKFDLDESRRKYIFNKTAEYQEDMDLKDCLKLQNKDAKIAVSIFYPNTLGLVK 721
DB 218 -WNNNAIKF-----QAWCDLLKQNNLY---GAGIVIIY--TMLPGFLK 254
QY 722 KL 723
DB 255 EL 256
RESULT 9
GGAA_BACSU STANDARD; PRT; 446 AA.
AC P46917;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Minor teichoic acids biosynthesis protein ggaA.
GN GGAA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Freymond P., Karamata D.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borrius R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maestl D., Nakai S., Noback M.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Oudega B., Park S.H.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara S., Prescott A.M.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Yamane K., Weitzneger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "the complete genome sequence of the Gram-positive bacterium Bacillus
subtilis.",
RL Nature 390:249-256(1997).
CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF GALACTOSAMINE-CONTAINING
CC MINOR TEICHOIC ACIDS.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U13979; AAA73512.1; -;
DR EMBL; Z99122; CAB15586.1; -;
DR Subtilist; BG11367; ggaA.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Transferase; Glycosyltransferase; Complete proteome.
SQ SEQUENCE 446 AA; 53148 MW; BB3698D9B6856C42 CRC64;
Query Match 3.8%; Score 191.5; DB 1; Length 446;
Best Local Similarity 21.5%; Pred. No. 0.0014;
Matches 99; Conservative 69; Mismatches 136; Indels 157; Gaps 21;
QY 434 LVSIYIPAYNCANYIQCVDLSALNOTV---VDLEVCINDGSDTNT---LEVINKLYGN 487
DB 1 MFSIIIPYNSENLYSIESVLNQSIGFKNIELILIDGSDVSPQICESFKNLYPNN 60
QY 488 PVRIMSKPNGGIASASNAVS--FAKYYIGOLDSDDDYLEDPAVE-----LCLKEFLKD 540
DB 61 --IKMKIENSGPSAARNCGLSNVSESKFIFGLDSDAFSQNALQSVYDFCDSEHVN 118
QY 541 KTLACVYTTNRVNPVPGSLIANGYNWPEFSREKLTTAMIAHFRMFTI---RAWLTDG 596
DB 119 AVLPVYITGEKE---GGHKLNNRF-----EAGT-----RVINLNDYKAIHYIG 160
QY 597 ---FNENIENAVDYDMFLKLSLVGKFKHLNKCYNRVLHGDNTSIKGLGKIQKHHFVV 652
DB 161 GTFYRHTLTSTVLFDESIFWE-----DAIFFNQL-----LKEKRYGAVA 202
QY 653 NQSLNRQGINVYNDKFDLDESRRKYIFNKTAEYQEDMDLKDCLKLQNKDAKIAVSIF 712
DB 203 -----EKGKRYK-----RKQDSLVDRSWFNKRYIYLLNECY 236
QY 713 PNTLNLGLVKLNLNIEYKNIFVILLHVDKNHLPD----- 748
DB 237 MTLMDSFNKYDVLPLQ--FLIVYH-KLFLPNYRDYVKSVLDQEQQRVVDFFIKV 293
QY 749 -----IKKEILAFYHKHQNVLNNDISYTSNRLIKTEAHLNINKLSQNLNLC-- 798
DB 294 LKFTDPQFIKEQDPMYKKEFMHLLKEN-----TEA-LENIKKERVLRHSSCTV 341
QY 799 -----EYIIFDHDLSLVKNDYSYAMMK 821
DB 342 TSAKIGRLRLTGHFNFQYEMKENDRIYK-----YFKR 377
RESULT 10
EX5B_BUCAI STANDARD; PRT; 1174 AA.
ID EX5B_BUCAI
AC P57529;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exodeoxyribonuclease V beta chain (EC 3.1.11.5).
GN RECB OR B0454.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbolic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS";
RL Nature 407:81-86(2000).
CC -!- FUNCTION: REQUIRED FOR EFFICIENT DNA REPAIR; IT CATALYZES THE
UNWINDING OF DOUBLE-STRANDED DNA AND THE CLEAVAGE OF SINGLE-
STRANDED DNA AND IT STIMULATES LOCAL GENETIC RECOMBINATION. ALL OF
THESE ACTIVITIES REQUIRE CONCOMITANT HYDROLYSIS OF ATP (BY
SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-
phosphooligonucleotides.
CC -!- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
(BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL; AF001119; BAB13152.1; -.
DR HSSP; P56255; 1QH8.
DR InterPro: IPR000212; UvrD-helicase.
DR Pfam: PF00580; UvrD-helicase; 1.
KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;
FT DNA repair; Complete proteome.
FT NP_BIND 25 32 ATP.
FT DOMAIN 342 347 POLY-LYS.
SQ SEQUENCE 1174 AA; 139037 MW; 668FE1336F1F190A CRC64;
Query Match 3.7%; Score 189; DB 1; Length 1174;
Best Local Similarity 18.6%; Pred. No. 0.0067;
Matches 200; Conservative 172; Mismatches 348; Indels 358; Gaps 53;
QY 14 NDVELALALFEKSAETVGRKIVFQIICKELSTNSVYSEDKNSVCDSLSLQTL 73
DB 250 NDELNKKIYSK-----FNLIKWNK-----ITEWAKSETKDTITP----- 285
QY 74 SNVKKLTSEKNSLKNKW-KSITGKKSENAEIRKVELVPKDFPKDLVLAPLDHVNDF 132
DB 286 SILKFTKKNIEKNTNNCTSKYVIFEESEKILKKFSL-----KNVLIYAVKIHQF 339
QY 133 TWYKNRKKSLGKIPVKNKNGLSIIPTFNRSRL-DITLACLNVQNTNTPFEVVDGGS 191
DB 340 LLKERKKKSL-----IGFNDLLSLLTKTKRKFLLDII-----KKYPAFI----- 382
QY 192 KENLLTIVQYEQKLDIKVYRQKDYQVQLCAVNRGLRTAKYDFVSILDCDMAPQLVH 251
DB 383 DEFQDIDIOQY-KIFNLKKNKNTVFL-----IGDPKQIAYSFRGAD-----IF 427
QY 252 SYLTLELNDIVLIGPRYVDTH-----NITAEQFLNDPYLESLPETATNNN 300
DB 428 SYLYAKS-----INKYVYLDTNWRSSINICKSINFLESQIN-PFIFKNTPIYI----- 475
QY 301 PSITSKGNISLDMRLEHFKKTDNLRLCDSPPRYFVAGNVAFS-----KEWLNKVGWDFEF 356

DB 476 PVPSSKNLKNFTINDVAQT-----PISFFLOEKEVSIDDYQVWISK--QCANEI 525
QY 357 NHWGGEDVEGYRLFAKGCFFRVIDGGMALHQBPPCKENETEREACKSLTKLVKEKVPY 416
DB 526 SFW-----LTCAKTGRAKITTKNGEKILUTANDIALIVR- 558
QY 417 IYRKLPIEDSHIRPLVSIYIPAYNC-----ANYIORCVDSAL---NOTVVD----- 462
DB 559 -NRKEADLIQDELEKLNIIISYSSNKSVPQTLDQAEILLWILEPENEILLQSQMAS 617
QY 463 -----LEVCIKNDG-STDNTEVINKLY-----GNNPRVRIM---SKPNGGTASAS 504
DB 618 HILKLSLVVENKNTISNKSNIYFIEIKLYEYHDIWEKIGVQMIKIMILEYQKNSFCTEIN 677
QY 505 NAAVSFAKGYITGOLSDDDYDPAVELCLKEFLK-----DKTLAC-----V 546
DB 678 ENHIKNLNLFLHJGELLQEQFQPHKKISLIRWFQKKISKTKQPEYNESIKCFDESPIKI 737
QY 547 YTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHFRMPTIRAH-----LTDGFENIE 602
DB 738 ITIHKSGLEYPIV-----WIPESID-----FKKSTLAIYHDQKSLKTFDENYS 782
QY 603 NAVDYDMFLKLSB-----VGRFKHLNKICYNR----- 629
DB 783 N-----KFLKIADDEERLAEDIRFLYVALTRSLHCSIGLACLIKKIKNRNNSDIHQSL 837
QY 630 --VLHGDN-----SIKKGIGQKNHFVV-----NOSLNR 658
DB 838 GFTIQGGKTNVEGLEKLLKLSI---NNFIEVKNNTDNFSRKPOTISLICKNFKLNK 894
QY 659 QGI-NYNYNDKFDLDESRRKYIFNKTAEOEEMDLKLIQNKDAK-----IAVSIFYP 713
DB 895 KNIRNTWSITSFSQLNK-----INKLSKHQKQVALKEL-CIKQEKKNQSLTIHFPK 948
QY 714 NTLNGL-----VKLANIIEYKNKNIPIVILHVDKNHLTPDIKEILAFYKHQVNLNN 768
DB 949 GKRTGLMIHYILKHLNHLNKNKNSNWFSCILEKYNIIK---WTSVLIYWKNIINTPLND 1005
QY 769 -----DISYTSNRLIKTEAH--LSNINKLS-----OLNLSN----- 797
DB 1006 EKILSRIDEKSSIRLEFFFPPIKNMLYSTELNKIQTOSINPTITSQPSLPNPKVGLMTG 1065
QY 798 -----CEYIFDNHDSLFVKNDY-----AYMKKYDVGMNFSALTHDTEKI 839
DB 1066 FVDLVFIWKKKYVILDYKSNLCKNNNFYSSIIHNKEIVKKRYHLYQO-----IYTI 1117
QY 840 NAHPPEKLIKITYFNND-----LRSMNVKASOGMFKYALPHELLTIKEVIT 889
DB 1118 AIHKYLQKLLKNYKDDFGVYFFFLRAIDCPKKNNGIF--YTMPN--YSLIKKTMT 1171
RESULT 11
EX5B_BORBU
ID EX5B_BORBU STANDARD; PRT; 1169 AA.
AC O51578;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exodeoxyribonuclease V beta chain (EC 3.1.11.5).
GN RECB OR B06033.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D., Hanson M.,
RA Peterson J., Karpavavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Wattley L., McDonald L., Artiach P., Bowman C.,


```
QY 67 TATQLLSNVKLTSESEKNSLKNKWSITGKKSNAEIRKVELVPKDPKDLVLAPL 126
Db 96 IETKNVILSLKKTASNSQSLPS-----KISKNIFFIKYPIILKKIHADKIL--- 144
QY 127 DHVNDFTYKRRKSL---GIKPVKNKIGLSIIPTNRSRILDITLACLVNQKNTNPF 182
Db 145 -----FTSPKVRVSLDVLGSLKLRNNI---IFSPTY-----INTIHVSSAKFNF-- 187
QY 183 EVVADDSKENLL---TIVOKYEOKLDIKVVRKDY-----GYOLCAVRNGLRLTA 231
Db 188 -----EKNILNKSTIIKNFNK---IKKIYSFSYSSNOTKKNFPL-----NIIYLSL 232
QY 232 K-----YDFVSILDCDMPAQQLVHVSXYLTLELDDNDIVLIGRPKYVDTH-----NITA 279
Db 233 KCKNTKTOFIDYKKNLQVEL-----QANIENNILQINKMK-VDSFLKMNISVG 279
QY 280 EQLNDPILIESLPTAT-----NNNPSITSKGNISDLWRLEHFKKTDN----- 323
Db 280 KVIFNNDYSISCVMSKTVIPSLYKNSINFOLKANFNVDHQL-IFKLSKDLNKMKNGL 338
QY 324 --LRLCDSFPRFV--AGNVA--FSKEWLNKVGWDFDEEFNHGGEDEFGYRFAKGC 376
Db 339 VFLNFSDFPF--FIKLSRNSLCVTKKNYIFKLKSF-----GVULKGI-----INNY 383
QY 377 FRVIDGMAIHQEP-----GKENETEREAGKSITLXIVREKVPYIYKLLPIED---- 426
Db 384 FFSKNIPTLODLPPIFIDIOGR-GDLNNIFLKKINFPIKQK--KFYKVIHPEDYIKY 440
QY 427 -----SHIHRIPLSYIIPAYNCANYIQRVDSALNQTVDLEVC 466
Db 441 NOYILKLGQINITGKSDRHTHYVHPKIDLY-----ANIMKK----- 478
QY 467 ICNDGSTDLTLEVINKLYGNN-----PRVRIMSKPN-----GGIASASNAVSPAKGY 515
Db 479 -----KLSILGALYKKNFNFTETGGINLLGKNKLYLRSGLKKNYISSI---Y 525
QY 516 IQLSDSDYLEPDAVELCLKEFLDKTLACVYTTNRNVP--DGLSIANGYNWPE--FSR 571
Db 526 ANNL--DYFFPK-----LOGRMOAKVNFNGNKPPIIISKILARDLNNNIYFKN 573
QY 572 EKLTTA-----MTAH-----HFRMTIRAWHLTDGFNENIENAVDYDMFLKSEVG 617
Db 574 IKVLGTGINNTFSKMLIYANKIHFKFYINTLHIQYTSNNHKQN---FSLLKSNRL- 629
QY 618 KFKHLNKI---CYNRVLHGDNTSIIKLGIO-----KKNHFVVVNNQSLNRQGINNY 666
Db 630 ---HNLINGAFNKTGHWGFFKINIRTFWQGVTAKKNNFIHYDHSNITNFIYQKS 686
QY 667 DKFDDLESRY-----IFNKT-AEYQEMDMKLKDLKIQNK----- 702
Db 687 IKKRNCFSSFLYNVKMSFNLNRSFISFESKLSINAKLKLGLKMSIDGAIFLKGNNTK 746
QY 703 -DAKLAVSIFYPN-----TLNGLVK-----KLNIIYNNKIF-----VILHVDK 742
Db 747 LEKKIKKIFIQIDFFKISMN-LIKNDPKFKWIIKKKKLNNKIFGLNIIDIYNNK 805
QY 743 NHLTDPDIKEILAEVAKOVNILLANDISYTSN-----RLIKT-----E 782
Db 806 N-----IKGEFI--FYKPPFP-----INFTTNEKVSQKFSQIKFGLTIYQPKVLAD 853
QY 783 AHLNLI-----NKLSQLNLCYEIIFDNHDSLFV-----KNDSD--- 815
Db 854 VHFKNIFIRSNILKAVITLFFPYFLGKVDNIKINOE-IMNKGNIILFTLKPFFKNNSADI 912
QY 816 ---YAYMKKYDVGMFNSALTHWDIEKINAHPPFKK 847
Db 913 ENNIASFNSKKSIVLFFPKIKVRFSSKLNHLHLYFSK 947

RESULT 13
Y025_MYCGE
ID Y025_MYCGE STANDARD; PRT; 298 AA.
AC P47271; Q49359;
```

```
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative glycosyl transferase MG025 (EC 2.4.1.10).
GN MG025.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37; PubMed=7569993;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 172-298 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing."
RL J. Bacteriol. 175:7918-7930(1993).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC STRONG, TO M.PNEUMONIAE MPN028.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U36682; AAC71241.1; -
DR EMBL: U02253; AAD12517.1; -
DR TIGR: MG025; -
DR InterPro: IPR001173; Glycos_transf_2.
DR Pfam: PF00535; Glycos_transf_2; 1.
KW Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
FT CONFLICT 200 201 EP -> DA (IN REF. 2).
SQ SEQUENCE 298 AA; 35044 MW; B7BCCC73EB2CA023 CRC64;

Query Match 3.5%; Score 176.5; DB 1; Length 298;
Best Local Similarity 21.8%; Pred. No. 0.0055;
Matches 80; Conservative 40; Mismatches 118; Indels 129; Gaps 14;

QY 434 LVSIIYIPAYNCANYIQRVDSAL--NOTVVDLEVCINDGSTDTNTLEVINKLYGNPRVR 491
Db 6 LFTVIPTNCCQYIKKALDLSLLQNEFLKQTVLVNDGSLDNTKEVSDYLIKYSNIS 65
QY 492 IMSKPNNGGTASASNAV--SFAKGYIIGQLSDSDYLEPDAVELCLKEFLDKTLACVYT- 548
Db 66 YFEKTNGNGSVINYYKKNKALGOYITVLDSDDYFLKDSFKKVRFFGHDMITGAFYCY 125
QY 549 TNRNVNPDGSLIANGYNWPEFSREKLTMTAMIAHFRMTIRAWHLTDGFNENIENAVDYD 608
Db 126 INEN-----KTRFLKPYFGK---TGVIKERTKLRTPHS----- 155
QY 609 MFLKLEVGKFKHLNKCYNRVLHGDNTSIKLGIOKKNHFVVVNNQSLNRQGINNYNDK 668
Db 156 -----QPIAKF-YSNKLFYE--LH----- 171
QY 669 FDDLDESRYIYNKTAETAEYQEEEMDMKLKDLKIQNKDAKIAVSIFY----- 712
Db 172 --DLKE--KLFFQDCLMYHDAINRVE-----SVFYREPLAVWFSTRPGNS 213
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2003, 02:13:54 ; Search time 121 Seconds
(without alignments)
1643.268 Million cell updates/sec

Title: US-09-842-484A-2
Perfect score: 5089
Sequence: 1 MNTLSQAIKAYNSNDYELAL.....SARKGENIPVKNKFIINSITL 965

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5089	100.0	965	2 Q9KJ99	Q9KJ99 pasteurella
2	4802	94.4	965	16 Q9CMP0	Q9CMP0 pasteurella
3	4791	94.1	965	2 Q9AHL6	Q9AHL6 pasteurella
4	4493.5	88.3	972	2 Q85457	Q85457 pasteurella
5	4486.5	88.2	972	2 Q68389	Q68389 pasteurella
6	294	5.8	604	16 Q921F9	Q921F9 rickettsia
7	292.5	5.7	327	16 Q8Y828	Q8Y828 listeria mo
8	292	5.7	337	16 Q9CLR9	Q9CLR9 pasteurella
9	287.5	5.6	350	2 Q9ALS8	Q9ALS8 campylobact
10	279.5	5.5	326	16 Q8XN37	Q8XN37 clostridium
11	276.5	5.4	706	2 Q88090	Q88090 enterococu
12	268.5	5.3	696	16 Q97P72	Q97P72 streptococ
13	265.5	5.2	348	16 Q8XN34	Q8XN34 clostridium
14	264.5	5.2	323	16 Q8YSL7	Q8YSL7 anabaena sp
15	261.5	5.1	303	16 Q9K6L6	Q9K6L6 bacillus ha
16	259.5	5.1	323	2 Q86893	Q86893 streptococ

17	259	5.1	342	16 Q8XN57	Q8XN57 clostridium
18	258.5	5.1	333	16 Q97H38	Q97H38 clostridium
19	257	5.1	316	2 Q06035	Q06035 lactococcus
20	256	5.0	321	16 Q8YSM2	Q8YSM2 anabaena sp
21	256	5.0	330	16 Q8XN54	Q8XN54 clostridium
22	254.5	5.0	324	16 Q8YSL1	Q8YSL1 anabaena sp
23	254	5.0	315	2 Q93TI5	Q93TI5 streptococ
24	253	5.0	321	2 Q9AFH3	Q9AFH3 streptococ
25	251	4.9	298	17 Q9U216	Q9U216 pyrococcus
26	249.5	4.9	283	2 Q93TI7	Q93TI7 streptococ
27	247	4.9	324	2 Q9AQJ0	Q9AQJ0 streptococ
28	246	4.8	322	2 Q9RG46	Q9RG46 streptococ
29	241.5	4.7	281	2 Q92GK2	Q92GK2 leptospira
30	241	4.7	324	2 Q8VLB0	Q8VLB0 streptococ
31	241	4.7	333	2 Q87182	Q87182 streptococ
32	241	4.7	333	16 Q8YW50	Q8YW50 anabaena sp
33	241	4.7	582	2 Q9AEE2	Q9AEE2 leptospira
34	241	4.7	582	2 Q8VTX6	Q8VTX6 leptospira
35	239	4.7	281	2 Q9S4F8	Q9S4F8 leptospira
36	239	4.7	340	2 Q87159	Q87159 vibrio chol
37	238.5	4.7	323	2 Q85000	Q85000 streptococ
38	238.5	4.7	324	2 Q86046	Q86046 streptococ
39	238	4.7	306	2 Q07339	Q07339 streptococ
40	238	4.7	316	16 Q8YSL6	Q8YSL6 anabaena sp
41	237.5	4.7	257	16 Q9K6R6	Q9K6R6 bacillus ha
42	237.5	4.7	306	2 Q9RP62	Q9RP62 escherichia
43	237	4.7	344	16 P71059	P71059 bacillus su
44	236.5	4.6	577	2 Q9S4F7	Q9S4F7 leptospira
45	236.5	4.6	1238	12 Q9EMP3	Q9EMP3 ansacta moo

ALIGNMENTS

RESULT 1

ID	Q9KJ99	PRELIMINARY;	PRT;	965 AA.
AC	Q9KJ99;			
DT	01-OCT-2000 (Tremblrel. 15, Created)			
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Chondroitin synthase CS.			
OS	Pasteurella multocida.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;			
OC	Pasteurella			
OX	NCBI_TaxID=747;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=P4679;			
RX	MEDLINE=20379058; PubMed=10818104;			
RA	DeAngelis P.L., Padgett-McCue A.J.;			
RT	"Identification and Molecular Cloning of a Chondroitin Synthase from			
RT	Pasteurella multocida Type F.;"			
RL	J. Biol. Chem. 275:24124-24129(2000).			
DR	EMBL; AF195517; AAF97500.1; -			
DR	HSSP; P39621; 10GO.			
DR	InterPro; IPR001173; Glycos_transf_2.			
DR	Pfam; PF00535; Glycos_transf_2; 2			
SQ	SEQUENCE 965 AA: 111392 MW; AFD55CD59D35C8C CRC64:			
Query Match 100.0%; Score 5089; DB 2; Length 965;				
Best Local Similarity 100.0%; Pred. No. 8.1e-256;				
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKRIIVEFQIIKCKEKLSTNSYVSEDKNSV 60			
DB	1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKRIIVEFQIIKCKEKLSTNSYVSEDKNSV 60			
QY	61 CDSSLDIATQLLSNVKLTLSSEKSNLKNWKSITGKSENAETRKVELVPKDFPKDL 120			
DB	61 CDSSLDIATQLLSNVKLTLSSEKSNLKNWKSITGKSENAETRKVELVPKDFPKDL 120			
QY	121 VLAPLPDHVNDFTWYKNRKKSLGKIPVKNKNIIGLSIIPTFNRSRILDITLACLVNQKTNV 180			

Db 121 VLAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIIPFNRSRLDITLACLVNQKTN 180
Qy 181 PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQICAVRNGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQICAVRNGLRTAKYDFVSILD 240
Qy 241 CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAQLNDPDLIESLPETATNN 300
Db 241 CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAQLNDPDLIESLPETATNN 300
Qy 301 PSITSKGNISLDWLEHFKKTDNLRLCDSPRYFVAGNVAESKELNKGWDFEENHWG 360
Db 301 PSITSKGNISLDWLEHFKKTDNLRLCDSPRYFVAGNVAESKELNKGWDFEENHWG 360
Qy 361 GEDVEFGYRLFAKCFRRVIDGGMALHQPPEPKENETEREAGKSITLKIYKVPYIYRK 420
Db 361 GEDVEFGYRLFAKCFRRVIDGGMALHQPPEPKENETEREAGKSITLKIYKVPYIYRK 420
Qy 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI 480
Db 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI 480
Qy 481 NKLYGNNPRVRIMSKPNGGSIASASNAAVSFAGYIYIGQLSDDDYLEDPAVELCLKEFLKD 540
Db 481 NKLYGNNPRVRIMSKPNGGSIASASNAAVSFAGYIYIGQLSDDDYLEDPAVELCLKEFLKD 540
Qy 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN 600
Db 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN 600
Qy 601 IENAVDYDMFLKSEVGKFKHLNKCINRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG 660
Db 601 IENAVDYDMFLKSEVGKFKHLNKCINRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG 660
Qy 661 INYNYDKFDDLDESRRKYIFNKTAQYQEMDMKLDKLIQNKDAKIAVSIYFNTLNGLV 720
Db 661 INYNYDKFDDLDESRRKYIFNKTAQYQEMDMKLDKLIQNKDAKIAVSIYFNTLNGLV 720
Qy 721 KKLNNIETYNKNTFVILHVDKNHLPDIIKKEILAFYHKHQVNILLNDISYTSNRLIK 780
Db 721 KKLNNIETYNKNTFVILHVDKNHLPDIIKKEILAFYHKHQVNILLNDISYTSNRLIK 780
Qy 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFNSALPHDWEIKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFNSALPHDWEIKIN 840
Qy 841 AHPPFKLLIKTYFNNDLRSMNVKASQGMFKYALPHELLIITKEVITSCOSIDSVPY 900
Db 841 AHPPFKLLIKTYFNNDLRSMNVKASQGMFKYALPHELLIITKEVITSCOSIDSVPY 900
Qy 901 NTEDIWFQFALLILEKKTGHVFNKTSITLYMPWERKLOWTNEQISAKKGENIPVNFKEII 960
Db 901 NTEDIWFQFALLILEKKTGHVFNKTSITLYMPWERKLOWTNEQISAKKGENIPVNFKEII 960
Qy 961 NSITL 965
Db 961 NSITL 965

RESULT 2
Q9CMP0 PRELIMINARY; PRT; 965 AA.
AC Q9CMP0;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Hypothetical protein PM0775.
GN .PM0775.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OC NCBI_TaxID=747;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006116; AAK02859.1; -
DR HSSP: P39621; LQGO.
DR InterPro: IPR001173; Glycos_transf_2.
DR Pfam: PF00535; Glycos_transf_2; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 965 AA; 111600 MW; 9C4B2CF80ELA6BD7 CRC64;
Query Match 94.4%; Score 4802; DB 16; Length 965;
Best Local Similarity 93.0%; Pred. No. 6e-241;
Matches 897; Conservative 42; Mismatches 26; Indels 0; Gaps 0;
Qy 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVEQIIEKCKEKLSTNSVSEDKNSV 60
Db 1 MNTLSQAIKAYNSNDYELALKLFKSAETYGKIVEQIIEKCKEKLSTNSVSEDKNSV 60
Qy 61 CDSISDITATQLLLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVPPDKDL 120
Db 61 CDSISDITATQLLLSNVKKLTLSSEKNSLKNKWSITGKSENAEIRKVELVPPDKDL 120
Qy 121 VLAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIIPFNRSRLDITLACLVNQKTN 180
Db 121 VLAPLPHVNDFTWYKRRKSLGKIPVKNKIGLSIIIPFNRSRLDITLACLVNQKTN 180
Qy 181 PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQICAVRNGLRTAKYDFVSILD 240
Db 181 PFEVVVADGSKENLLTIVQYEQKLDIKYVRQKDYGYQICAVRNGLRTAKYDFVSILD 240
Qy 241 CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAQLNDPDLIESLPETATNN 300
Db 241 CDMAPOQLWHSYLTLELNDIVLIGPRKYVDTHNTAQLNDPDLIESLPETATNN 300
Qy 301 PSITSKGNISLDWLEHFKKTDNLRLCDSPRYFVAGNVAESKELNKGWDFEENHWG 360
Db 301 PSITSKGNISLDWLEHFKKTDNLRLCDSPRYFVAGNVAESKELNKGWDFEENHWG 360
Qy 361 GEDVEFGYRLFAKCFRRVIDGGMALHQPPEPKENETEREAGKSITLKIYKVPYIYRK 420
Db 361 GEDVEFGYRLFAKCFRRVIDGGMALHQPPEPKENETEREAGKSITLKIYKVPYIYRK 420
Qy 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI 480
Db 421 LLPIEDSHIHRIPLVSIYIPAYNCANYIQRCDVSALNQTVDLEVCICNDGSTNTLEVI 480
Qy 481 NKLYGNNPRVRIMSKPNGGSIASASNAAVSFAGYIYIGQLSDDDYLEDPAVELCLKEFLKD 540
Db 481 NKLYGNNPRVRIMSKPNGGSIASASNAAVSFAGYIYIGQLSDDDYLEDPAVELCLKEFLKD 540
Qy 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN 600
Db 541 KTLACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHLTDGFNEN 600
Qy 601 IENAVDYDMFLKSEVGKFKHLNKCINRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG 660
Db 601 IENAVDYDMFLKSEVGKFKHLNKCINRVLHGDNTSIKLGIOKKHFFVNVQSLNRQG 660
Qy 661 INYNYDKFDDLDESRRKYIFNKTAQYQEMDMKLDKLIQNKDAKIAVSIYFNTLNGLV 720
Db 661 VSNYNTDEFNDLDESRRKYIFNKTDYQEEIDILKQIKVQRDAKVAISIFYPNRLDGLV 720
Qy 721 KKLNNIETYNKNTFVILHVDKNHLPDIIKKEILAFYHKHQVNILLNDISYTSNRLIK 780
Db 721 KKLNNIETYNKNTFVILHVDKNHLPDIIKKEILAFYHKHQVNILLNDISYTSNRLIK 780
Qy 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFNSALPHDWEIKIN 840
Db 781 TEAHLNSINKLSQNLNCEYIIFDNHDSLFVKNDSYAYMKKYDVGWGFNSALPHDWEIKIN 840

Best Local Similarity 86.8%; Pred. No. 5.9e-225;		Matches 844; Conservative 61; Mismatches 60; Indels 7; Gaps 2;	
Qy	1	MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVFQIKCKEKL---STNS---YVS	53
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVFQIKCKEKL---STNS---YVS	60
Qy	54	EDKNSVCDSDIATQALLSNVKKLTSESEKNSLKNKWSITGKSENAEIRKVELVP	113
Db	61	KEEVNVCDSPLDIATQALLSNVKKLVSDSEKNTLKNKWLTEKSENAEIRAVALVP	120
Qy	114	KDFPKDLVLAHPDHNDFTWYKNNKKSGLGKPVKNKLGSLIIPTFNRSILDTLACL	173
Db	121	KDFPKDLVLAHPDHNDFTWYKNNKKSGLGKPVKNKLGSLIIPTFNRSILDTLACL	180
Qy	174	VNOKTNPFEVVDGSGENLTTVQYEOKLDIKYVRQDYGQYQCAVRNGLRTAKY	233
Db	181	VNOKTHYPFVIVTDGSGEDLSPIRQYENKLDIYVRQDNGFQASARNMGLRLAKY	240
Qy	234	DFVSILDCDMPAQLWVHSYITELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLP	293
Db	241	DFIGLLDCDMPAPNLWVHSYVAELLEDLTIIGPRKYVDTHNITAEQFLNDPYLESLP	300
Qy	294	ETATNNPISITSGNLSLDWRLEHFKKTDLNRLCDSPPRYFVAGNVAFSKWLNKVGWFD	353
Db	301	EYKTNVAAKGEVTVSLDWRLEQFEKTENLRLSDSPFFFAAGNVAFKAKWLNKSGPFD	360
Qy	354	EEFNHWGGEDEVFGYRLFAGCGFFRVIDGGMALHQBPGPKENETEREAGKSTLKIYKEK	413
Db	361	EEFNHWGGEDEVFGYRLFAGCGFFRVIDGGMALHQBPGPKENETEREAGKSTLKIYKEK	420
Qy	414	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQVVDLEVCICNDGST	473
Db	421	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQVVDLEVCICNDGST	480
Qy	474	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC	533
Db	481	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC	540
Qy	534	LKEFLKDKTLACVYTTNRNVPDGLSIANGYNWPFESREKLTAMTAHFRMETIRAWHL	593
Db	541	LKEFLKDKTLACVYTTNRNVPDGLSIANGYNWPFESREKLTAMTAHFRMETIRAWHL	600
Qy	594	TGCFNENIENAVDYMFLKLVSEVGFKNLKNKICYNRVLHGDNTSIRKLGIOKKNHFVYN	653
Db	601	TGCFNENIENAVDYMFLKLVSEVGFKNLKNKICYNRVLHGDNTSIRKLGIOKKNHFVYN	660
Qy	654	QSLNRQGINVYNDKFDLDESRYIPNKTAEQEEMDMKDLKIQNKDAKTAVSTFYP	713
Db	661	QSLNRQGITVYNDKFDLDESRYIPNKTAEQEEDILDKIKIQNKDAKTAVSTFYP	720
Qy	714	NTLNGLVKKLNITENYKNTFVILHVDKNHLPDIIKKEILAFYKHQVAILLNDISYY	773
Db	721	NTLNGLVKKLNITENYKNTFVILHVDKNHLPDIIKKEILAFYKHQVAILLNDISYY	780
Qy	774	TSNRLIKTEAHLNKLNSQLNCEYIIFDNHSLFVNKDSYAMKYDVGMMFSAETH	833
Db	781	TSNRLIKTEAHLNKLNSQLNCEYIIFDNHSLFVNKDSYAMKYDVGMMFSAETH	840
Qy	834	DWIEKINHPFKLITKYTFNDNDRSNVKGASQGMFKYALPHELLTIKEVITSQOS	893
Db	841	DWIEKINHPFKLITKYTFNDNDRSNVKGASQGMFKYALPHELLTIKEVITSQOS	900
Qy	894	IDSPYEYNTEDIWFQFALLILEKKTGHVFNKSTLITYMPWRKQLQWNTQOSAKKGNI	953
Db	901	IDSPYEYNTEDIWFQFALLILEKKTGHVFNKSTLITYMPWRKQLQWNTQOSAKKGNI	960
Qy	954	PVNKEFIINSITL 965	
Db	961	PVNKEFIINSITL 972	
RESULT 5			

068389	PRELIMINARY;	PRT;	972 AA.
ID	O68389		
AC	O68389;		
DT	01-AUG-1998 (TrEMBLrel. 07, Created)		
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Hyaluronan synthase.		
GN	PHAS OR HAS.		
OS	Pasteurella multocida.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Pasteurella.		
OX	NCBI_TaxID=747;		
EN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=P-1059, AND ATCC 15742;		
RX	MEDLINE=98192645; PubMed=9525958;		
RA	DeAngelis P.L., Jing W., Drake R.R., Achyuthan A.M.;		
RT	"Identification and molecular cloning of a unique hyaluronan synthase		
RT	from Pasteurella multocida.";		
RL	J. Biol. Chem. 273:8454-8458(1998).		
EN	[2]		
RC	SEQUENCE FROM N.A.		
RA	Fuller T.E., Kennedy M.J., Lowery D.E.;		
RT	"Identification of Pasteurella multocida virulence genes in a		
RT	septicemic mouse model using signature-tagged mutagenesis.";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF036004; AAC38318.1; -		
DR	EMBL; AF237926; AAF68412.1; -		
DR	HSP; P39621; 10G0.		
DR	InterPro; IPR001173; Glycos transf. 2.		
DR	Prfam; PF00535; Glycos transf. 2; 2		
SQ	SEQUENCE 972 AA; 111838 MW; 7D65D024C41ED362 CRC64;		
Query Match 88.2%; Score 4486.5; DB 2; Length 972;			
Best Local Similarity 86.6%; Pred. No. 1.4e-224;			
Matches 842; Conservative 63; Mismatches 60; Indels 7; Gaps 2;			
Qy	1	MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVFQIKCKEKL---STNS---YVS	53
Db	1	MNTLSQAIKAYNSNDYELALKLFKSAETVGRKIVFQIKCKEKL---STNS---YVS	60
Qy	54	EDKNSVCDSDIATQALLSNVKKLTSESEKNSLKNKWSITGKSENAEIRKVELVP	113
Db	61	KEEVNVCDSPLDIATQALLSNVKKLVSDSEKNTLKNKWLTEKSENAEIRAVALVP	120
Qy	114	KDFPKDLVLAHPDHNDFTWYKNNKKSGLGKPVKNKLGSLIIPTFNRSILDTLACL	173
Db	121	KDFPKDLVLAHPDHNDFTWYKNNKKSGLGKPVKNKLGSLIIPTFNRSILDTLACL	180
Qy	174	VNOKTNPFEVVDGSGENLTTVQYEOKLDIKYVRQDYGQYQCAVRNGLRTAKY	233
Db	181	VNOKTHYPFVIVTDGSGEDLSPIRQYENKLDIYVRQDNGFQASARNMGLRLAKY	240
Qy	234	DFVSILDCDMPAQLWVHSYITELLEDNDIVLIGPRKYVDTHNITAEQFLNDPYLESLP	293
Db	241	DFIGLLDCDMPAPNLWVHSYVAELLEDLTIIGPRKYVDTHNITAEQFLNDPYLESLP	300
Qy	294	ETATNNPISITSGNLSLDWRLEHFKKTDLNRLCDSPPRYFVAGNVAFSKWLNKVGWFD	353
Db	301	EYKTNVAAKGEVTVSLDWRLEQFEKTENLRLSDSPFFFAAGNVAFKAKWLNKSGPFD	360
Qy	354	EEFNHWGGEDEVFGYRLFAGCGFFRVIDGGMALHQBPGPKENETEREAGKSTLKIYKEK	413
Db	361	EEFNHWGGEDEVFGYRLFAGCGFFRVIDGGMALHQBPGPKENETEREAGKSTLKIYKEK	420
Qy	414	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQVVDLEVCICNDGST	473
Db	421	VPIYRKLLPIEDSHIRPLVSIYIPAYNCANYIQRQVDSALNQVVDLEVCICNDGST	480
Qy	474	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC	533
Db	481	DNTLEVINKLYGNPRVRIMSKPNGGIASASNAAYSFAGYYIGQLSDSDYLEPDAVELC	540

```
Qy 534 LKEFLKDKTKTACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHL 593
Db 541 LKEFLKDKTKTACVYTTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMFTIRAWHL 600
Qy 594 TDGNEENIENAVDYMFLKSEVGKFKHLNKCICYNRVLHGDNSTSIKLGIOKKHNFVVVN 653
Db 601 TDGFNEKIENAVDYMFLKSEVGKFKHLNKCICYNRVLHGDNSTSIKLGIOKKHNFVVVN 660
Qy 654 QSLNRQGINYNVDKFDLDESRYIFNKTAEOYOEEMDKDLKLIQNKDAKIAVISIYP 713
Db 661 QSLNRQGIYNYDEFDLDESRYIFNKTAEOYOEEDILDKILQNKDAKIAVISIYP 720
Qy 714 NTLNGLVKKLNIIYEYKNKFIIVIIHLVDKNHLLTPDIKEILAFYKHQHVNILLNNDISY 773
Db 721 NTLNGLVKKLNIIYEYKNKFIIVIIHLVDKNHLLTPDIKEILAFYKHQHVNILLNNDISY 780
Qy 774 TSNRLIKTEAHLNINLKLSQNLNCEYIIFDNHDSLFYKNDISYAYMKKYDYGMPFSALTH 833
Db 781 TSNRLIKTEAHLNINLKLSQNLNCEYIIFDNHDSLFYKNDISYAYMKKYDYGMPFSALTH 840
Qy 834 DWIEKINARPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIKEVITSQS 893
Db 841 DWIEKINARPPFKLIKTYFNNDLRSNMVKGASQGMFKYALPHELLTIKEVITSQS 900
Qy 894 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSITLTYPMPWERKLTQWNEQIOSAKKGNI 953
Db 901 IDSVPYNTEDIWFQFALLILEKKTGHVFNKTSITLTYPMPWERKLTQWNEQIOSAKKGNI 960
Qy 954 PVNKFINSITL 965
Db 961 PVNKFINSITL 972
```

```
RESULT 6
Q92IF9 PRELIMINARY; PRT; 604 AA.
ID Q92IF9;
AC Q92IF9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative two-domain glycosyltransferase.
GN RC0461.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AE008609; AAL02999.1; -.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 2.
KW Transferase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 604 AA; 69121 MW; 514CDD81DC3FB3AB CRC64;
```

Query Match 5.8%; Score 294; DB 16; Length 604;
Best Local Similarity 24.8%; Pred. No. 1.3e-07;
Matches 124; Conservative 71; Mismatches 213; Indels 92; Gaps 19;

```
Qy 419 RKLPIEDSHIRIPLVSIYIPAYNCANYIQCVDLSALNQTVVDLEVCICNDGSDTN-TL 477
Db 4 RKDKPLNIYH----TLVSIIPVYNGANYKMEAINSAQAQYKNEIIIVNDGSKDNGET 59
Qy 478 EVINKLYGNPRVRIMSPNGIASAANAVSFAKGYIGQLDSDDDYLEPDAVE----LC 533
Db 60 ERVALSYGD--KIRFYFKNGCGGSALNGYIKNMQGFYSWLSHDDIYYPNKHIEQVDDIL 117
```

```
Qy 534 LKEFLKDKTKTACVY-----TTNRNVNPDGSLIANGYNWPEFSREKLTAMIAHHFRMF 586
Db 118 NKLDNKDFTIYGGYELIDEKGNLSRYIKPDSVLPINKLN---ISLLPLRLGLIHGCSLLM 174
Qy 587 TIRAWHLTDGNEENIENAVDYMFLKSEVGKFKHLNKCICYNRVLHGDNSTSIKLGIOKK 646
Db 175 PAKYFHEVGIFNEALPTTQDYDLWFKIFRVAIPHEDSILIKSRFHSQGS--KITSNHNE 233
Qy 647 N-----HFVVVYNQSLNRQGINYNVDKFDLDESRYIFNKTAEOYOEEMDKDLK 697
Db 234 ECNVWSSFLHELTEEMIKMEGSPYL-----FLTRATFLSNNT-PYKKAACDLANTWA 286
Qy 698 LIQNKDAKIAVISIYPNTLNGLVKKLNII-IEYNKNIIFVILH----VDKNHLLTPDIKE 752
Db 287 KOVLNDTRISVIIPVYNINNAIEAKSVLIQTHKNFEILIIDGSTDIDSELTAICKRD 346
Qy 753 -ILAFYKHQHVNILLNNDISYITSNRLIKTEAHLNINLKLSQNLNCEYIIFDNHDSLFV 811
Db 347 KRIKYFHK-----KNEGPAARN-LGIKNAIGKYIAFLDSDDLFY 385
Qy 812 KNDISYAYMKKYDVGMPFSALTHDWIEKINARPPFKLIKTYFNNDLRSNMVKGASQGMF 871
Db 386 K-DKIEIQLKWEENNF-IFSHTSYHKINE-----KCKYIESVHSGLF 426
Qy 872 MKYALPHELLTIKEVITSC 891
Db 427 SGNVFP-----QVIQTC 438
```

```
RESULT 7
Q8Y828 PRELIMINARY; PRT; 327 AA.
ID Q8Y828;
AC Q8Y828;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein lmo1090.
GN LMO1090.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-E / SEROVAR 1/2A;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitouarn A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC99168.1; -.
DR ListiList; LMO1090; -.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 327 AA; 37843 MW; F7480656043EC2FC CRC64;
```

Query Match 5.7%; Score 292.5; DB 16; Length 327;
Best Local Similarity 27.5%; Pred. No. 7.7e-08;
Matches 98; Conservative 44; Mismatches 133; Indels 81; Gaps 9;

```
Qy 433 PLVSIYIPAYNCANYIQCVDLSALNQTVVDLEVCICNDGSDTNLEVINKLYGNPRVRI 492
Db 3 PLVSIIPVYNGANYKVKRCLDSVLEQTYHNLEIVVNDGATDSAKVIKISDN--RIRY 60
```

Qy	493	MSKPNGGIASNAAFVAFKGYIGQLSDSDYLEPDPAVELCLKEFLKDKTKLACVYTTNRN	552
Db	61	FEKENGQATARNFGLDVGDIYVMVDSDDYISKNIVETCL-DTVQKTNADLVFTSYN	119
Qy	553	VNPQGS-----LIANGYNWPEFSREKLTATTAMIAHHFPMFTIRAWHLTDGPF---	597
Db	120	VNQGKMQYIKRDKGIKVLDAGPTPNKFKYQADLWKG-----SRFPVGYWYEDLGIIPV	173
Qy	598	-----NENTENAVDY---DMFLKLSEVGKFKHLNKICYNRVLHGDNSTIKKLGIOKKN	647
Db	174	VTLKAKNPVKIQDALYYITDRADSQSNIOQVDHFLDVI---MLENVEYTELKKLGI---	227
Qy	648	HFVVVQSLNRQGINYYNDKFDLDESRYIFNKTAIEYQEEMDKDL-----KLIONKDA	704
Db	228	-----YEEKDQLAYLYIEHLIYRLV	249
Qy	705	KTAVSFYPTNLGLVKKLNIIIEYNKNIFVLIILHVDKNHLPDIIKKEILAFYHKH	760
Db	250	RKAIYITNKQERKKLIKISTIOEKFPNWSYPYQAGGKLTATLKKRALWLYLHH	305
RESULT 8			
ID	Q9CLR9	PRELIMINARY;	PRT; 337 AA.
AC	Q9CLR9		
DT	01-JUN-2001 (TremBLrel. 17, Created)		
DT	01-JUN-2001 (TremBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TremBLrel. 19, Last annotation update)		
DE	Hypothetical protein PM1140.		
GN	PM1140.		
OS	Pasteurella multocida.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Pasteurella.		
OX	NCBI_TaxID=747;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PM70;		
RX	MEDLINE=21145866; PubMed=11248100;		
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;		
RT	"Complete genomic sequence of Pasteurella multocida PM70.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).		
DR	EMBL: A5006155; AAK03224.1; -		
DR	InterPro: IPR001173; Glycos.transf_2.		
DR	Pfam: PF00535; Glycos.transf.2; 1.		
DR	Hypothetical protein; Complete proteome.		
SK	SEQUENCE 337 AA; 39267 MW; 8639CBF5F700DB4 CRC64;		
Query Match 5.7%; Score 292; DB 16; Length 337;			
Best Local Similarity 27.0%; Pred. No. 8.5e-08;			
Matches 96; Conservative 67; Mismatches 124; Indels 68; Gaps 16;			
Qy	433	PLVSIYPANCANYIQRCDVSALNQTVVDLEVCICNDGSTNTLEVINKLYGNPRVRI	492
Db	7	PLVSLICAYNADRYIEECIDALLNQYKNLEIVVVDGSDDTLSKLHYPAGKDPRIKI	66
Qy	493	M-SKPNGGIASNAAFVAFKGYIGQLSDSDYLEPDPAVELCLKEFLKDKTKLACV---	549
Db	67	INNEENKGFASLNIGIASINGDYLDARTDADITKPEWIEKILGYMLSHPOIIAMGSYLT	126
Qy	550	NRNVNPDGSLIANGY-----NWPE-FSRKLTAMIAHH-----FRMFTIRAW	591
Db	127	ILSEGDGNSLIANYEHGDEWRNPLSHREIVEAMLFRNPIINNSMIVKSTVFRHGLR--	184
Qy	592	HLTDGFENIENAVDYMFLKLSEVGKF-KHLNKICYNRVLHGDNSTIKKLGIOKKNHFV	650
Db	185	-----FDPAYQHTEDYQFWEVSRGELANYPESLVYVR-LH--NTQTSSLHKNYQN---	233
Qy	651	VVNQSLNRQGINYYNDKFDLDESRYIFNKTAIEYQEEMDKDLKLIONKDAKI	706
Db	234	LMAKKIRKRAINYY----LQDLGVIRHLEDIFFHDIETIQAE---LASLSLDN---CI	283
Qy	707	AVSIFYPNTNLGLVKKLNIIIEYNKNIFVLIILHVDKNHLPDIIKKEILAFYHKH	761
Query Match 5.6%; Score 287.5; DB 2; Length 350;			
Best Local Similarity 25.5%; Pred. No. 1.5e-07;			
Matches 105; Conservative 66; Mismatches 129; Indels 111; Gaps 16;			
Qy	433	PLVSIYPANCANYIQRCDVSALNQTVVDLEVCICNDGSTNTLEVINKLYGNPRVRI	492
Db	5	PLVSIIPCYNAENFTENCINSITNGYINYEIICVDGSDTNTLKILNLSINNSRLKA	64
Qy	493	MSKPNGGIAS-ASNAAFVAFKGYIGQLSDSDYLEPDPAVELCLKEFLKDKTKLACVY---	548
Db	65	YSINHTGIPSVVKVNYGLKAKGEFLILSDDDMITTEYFLEKGIKYPQDNPVDIILYPIKF	124
Qy	549	--TNRVNPDGSLIANGYNWPEFSREKLTATTAMIAHHFPMFTIRAWHLTDGFENIENAVD	606
Db	125	MFSNNYKIIGGIYNNSLNISDVNLGATNKIISGR-----DAFRNIYNNKLI	172
Qy	607	YDMFLKLSEVGKFKHLNKICYNRVLHGDNSTIKKLGIOKK-----NHFVVVQSLNRQ	660
Db	173	GFPEYK-KTIDKIINFNEESFN-----GDEYSFREHLLQAKKIAFIIDTEFVYVN--FNQ	225
Qy	661	I-----NYNYDKFDLDESRYIFNKTAIEYQEEMDKDLKLIONKDAKIA	707
Db	226	ITKKIGVHHWDTKWTFNLEKL-----AQHNVEK-----KLIK-KINKIR	265
Qy	708	VSIFYPNTNLGLVKKLNIIIEYNKNIFVLIILHVDKNHLPDIIKKEILAFYHKHVNILLN	767
Db	266	YSIYV-----ELCIKFNKTEYL-----FSQNEKNIL-292	
Qy	768	NDISYVTSNRLLKTEAHLNLSINKLSQ-LNLNC-----EYIFDNHDSLFVK	812
Db	293	-----NKILEKNHLSRINSIFDFLFYCKKDEKGYIKFYKITYFYK	335
RESULT 10			
ID	Q8XN37	PRELIMINARY;	PRT; 326 AA.
AC	Q8XN37		
DT	01-MAR-2002 (TremBLrel. 20, Created)		
DT	01-MAR-2002 (TremBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TremBLrel. 21, Last annotation update)		
DE	Capsular polysaccharide biosynthesis protein.		
GN	CP50501.		
OS	Clostridium perfringens.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;		

```
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003187; BAB80207.1; -.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Complete proteome.
SQ SEQUENCE 326 AA; 38553 MW; 76503BF0B5DE1C03 CRC64;

Query Match          5.5%; Score 279.5; DB 16; Length 326;
Best Local Similarity 25.2%; Pred. No. 3.6e-07;
Matches 92; Conservative 66; Mismatches 114; Indels 93; Gaps 13;

Qy 435 VSIIPAYNCANYIQCVDLSALNQTVDVLEVCICNDGSDTNTLEVINKLYGNPNRVRIMS 494
    ||| ||| ||| :||:||||: ||: ||||| :||| :|||
Db 4 VSIIPINVENYLDKCLDTAVNQTGLGCIETIGIDGSDTKSRDLNKYAKKYSFIKIN 63
    ||| ||| ||| :||| ||| ||||| ||: ||| :|||
Qy 495 KPNGGIASASNAVSFAKGYIGOLSDDDYLEPDAVELCLKEFLDKDKTLACVYTTNRVN 554
    ||| ||| ||| :||| ||| ||||| ||: ||| :|||
Db 64 KENGLSSARTGTINAKGEYIFELSDDDYIEYDSMKICYELATYKNDLIVSFDRESFD 123
    ||| ||| ||| :||| ||| ||||| ||: ||| :|||
Qy 555 P-----DGSIIANGYNMPEFSREKLTMTAM----- 578
    ||| :| :| :| :| :| :|
Db 124 GEGFTKLLKYDRKDILDGNKMGCEYCELIKKNKYTASACLNFRREFLDINLNFYD 183
    ||| :| :| :| :| :| :|
Qy 579 -IAHFRMFTTRAWHLTDGFNENENAVDYDMFLK-----LSEVGKFKHLNKC 626
    || :||: ||| :| :| :| :| :| :|
Db 184 GILHEDELYTVKATLA-----NKVRYIPYKLYFRRLRSNSITQSISE---KHL---- 230
    ||| :| :| :| :| :| :| :| :| :|
Qy 627 YNRVLHGNTSIKKLGIOKKNHVVVNSLNROGINYN---YDKFDDLDSESKYIENKT 683
    ||| :| :| :| :| :| :| :| :| :|
Db 231 YGRLLAARESVHYLYGKD-----VFNDNTWCLLNHISNMYFMKDOIYENKR-IYNK- 283
    ||| :| :| :| :| :| :| :| :| :|
Qy 684 AYEQEMDMKDLKLQIONKDAKIAVSIFYPNTLNGVLKVLKNNIIEYNKNIFFIILHVDKN 743
    ||| :| :| :| :| :| :| :| :| :|
Db 284 -RYIEDKIINDIK--SKKD-----IKSFKLDLQYR--PFLIITYTIKK 321
    ||| :| :| :| :| :| :| :| :| :|
Qy 744 HLTPD 748
    :| |
Db 322 YLKKD 326

RESULT 11
O88090 PRELIMINARY; PRT; 706 AA.
AC O88090;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative glycosyl transferase (Fragment).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OGLRF;
RX MEDLINE=98380380; PubMed=9712783;
RA Xu Y., Murray B.E., Weinstock G.M.;
RT "A cluster of genes involved in polysaccharide biosynthesis from
RT Enterococcus faecalis OGLRF.";
RL Infect. Immun. 66:4313-4323(1998).
DR EMBL; AF071085; AAC35930.1; -.
DR InterPro; IPR001173; Glycos_transf_2.
DR InterPro; IPR000655; HTH_Cro.
```

```
DR Pfam; PF00535; Glycos_transf_2; 2.
KW PRINTS; PR00030; HTHCRO.
KW Transferase.
FT NON_TER 706 706
SQ SEQUENCE 706 AA; 80401 MW; 8710AD60C7D36315 CRC64;

Query Match          5.4%; Score 276.5; DB 2; Length 706;
Best Local Similarity 21.5%; Pred. No. 1.3e-06;
Matches 151; Conservative 115; Mismatches 268; Indels 167; Gaps 25;

Qy 260 DNDIVLIGPRKYVD--THNITAEQFLNDPLYLIESLPETATNNNPSTSGNLSLDWR--- 314
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 2 NEDIKVIIFSIVRDKTTNNLTITGHALDITITCS--PTFTINNNOVSAYINQVRLRDEV 59
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 315 -----LEHFKKTDNLRCLDSPERYFVAGNVAFSEKWLNK-----VGWF 352
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 60 NOIYTEPAIEAGFVVTLEGIKOKKVL-----PFHQSSAHVYTVDFPLNKKYPIVPGTE 114
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 353 DEEFNHGGEDVEFGYRLFAGKCFRVIDGGMAIHQEPFGKENETEREAGKSITLKVKE 412
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 115 DKVTRLM-----IKAKGFKYMAKNGIS-HTIQRAKIEKLNRQA--SYPNWLARN 161
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 413 KVPYTYRKLPLEDSHIHRIPLVSIYIPAYNC-ANVIQRCVDSALNQTVDVLEVCICNDG 471
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 162 EVLDI--EAMTOEIAITFHYQPKISIAMPVYVVEKWLRLCIDSILNQVYTNWELCMADDA 219
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 472 STD-NTLEVINKLYGNPNVRIM-SKPNGGIASASNAVSFAKGYIGOLSDDDYLEPDA 529
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 220 STDPNKKILTQQLDERIRVVFREQNGHISEATNSALAIATGEFVALLDNDDELAINA 279
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 530 VELCLKEFLKDKTLACVYTTNRNPNVDPGSLIANGYNMPEFSREKLTMTAMIAHFRMFTIR 589
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 280 FVEVVKVLNENPELDLIYSDEDKIDMDGNRSDPAFK-PDMSPDLLLTGTVYISHLGVYRRS 338
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 590 AWHLDGFGNENENAVDYDMFLKLSB---VGKFKHLNKCYNRVLHGDNTSTKK--LGIQ 644
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 339 ILEEIGFKRGKGEQSDYDLVLRFTKERTHTHPVLYVYWRMLPTSTAVDQGSKGVA 398
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 645 KKNHFVVVNSLNROGINYN-----YDKFDDLDSESKYIFNKTAEQEEMDMKDL 696
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 399 FEAGLRVAVODALVRGINGHATHGAANGUDYVDYDI-ESEK----- 438
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 697 KLIONKDAKIAVSIFYPNTLNG---LVKKNLNIIEYNKNIFFIILHVDKNHLTPDIKKEI 753
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 439 -----LVSIIP-TKNGYKDVQRCVSSIETTYQNYEIMADNGSDTPDKMHELY 487
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 754 LAFYHKHQVNILLNN-DISYYSNRLIKTEAHLNKLNSQLNKLNCEYIIFDNHDSLFVK 812
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 488 AFEEOQLPGRFFVESIDIPF-----NFTINNRAAKKAHGEYLLFLNNDTEVIT 536
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 813 ND-----SYA-----YMKK 821
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 537 ENWILMVSFAQERIGCVGAKLLYPNNTVQHVAGVILGLGGVAGHGHYGPBGDLGYFGR 596
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Qy 822 YDVGMNFSALTDHWI----EKINAHPPFKKLTKTYFNDNDL 858
    ||| ||| ||| :||: ||| :| :| :| :| :| :|
Db 597 LAINVYSAVTAACLLMKKADEDAVGGFEEAFTVAFNDVDL 637
    ||| ||| ||| :||: ||| :| :| :| :| :| :|

RESULT 12
Q97P72
ID Q97P72 PRELIMINARY; PRT; 696 AA.
AC Q97P72;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycosyl transferase, family 2/glycosyl transferase family
DE 8.
GN SPI771.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
```

```

RN      SEQUENCE FROM N.A.
RP      STRAIN=TIGR4;
RC      MEDLINE=21357209; PubMed=11463916;
RA      Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA      Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA      Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA      Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA      McDonalld L.A., Feldblyum T.V., Angiuoli S., Dickenson T., Hickey E.K.,
RA      Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., C.M.;
RA      Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT      "Complete genome sequence of a virulent isolate of Streptococcus
RT      pneumoniae";
RL      Science 293:498-506(2001).
DR      EMBL; AE007470; AAK75845.1;
DR      TIGR; SPI771;
DR      InterPro; IPR001173; Glycos_transf_2.
DR      pfam; PF00535; Glycos_transf_2; 1.
DR      pfam; PF01501; Glyco_transf_8; 1.
DR      Transferrase; Complete proteome.
KW      SEQUENCE 696 AA; 81336 MW; ADCA9C01AAB32C7 CRC64;

Query Match      5.3%; Score 268.5; DB 16; Length 696;
Best Local Similarity 20.5%; Pred. No. 3.3e-06;
Matches 101; Conservative 96; Mismatches 181; Indels 115; Gaps 17;

QY 435 VSIYIPAYNCANYIQCVCDSALNQTVWVLEVCICNDGSTDNTLEVINKLYGNPRVRIMS 494
DB 7 ITIVFVNVNENYLRKCLDSITQYKNIIEIVVNDGSTDASGEICKEPSEMDHRILEYE 66
QY 495 KPNGGIASASNAASFAGYIYGQDSDYDLEPDVAVELCLK---FELDKTLACVYTTNR 551
DB 67 QENAGLSAARNTGLNMSGNYTTFVDSDDIEQDYETLYKKIVEYQADIAVGNYSFNE 126
QY 552 N-----VNPDSGLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHL-- 593
DB 127 SEGMYFHILGDSYKVEKVDNVSIFENLYE---TQEMKSFALISAWGLKARLPEQLR 182
QY 594 -----TDGFNENIENAVDYDMFLKSEVGKFKHLNKICY-NRVLUHGD----- 634
DB 183 FDIGKLGEDGY-----LNQKVL-LSE--KVLYLNKSLYAIRIRKGLSLRSVWTEKWMH 232
QY 635 -----NTSKKLGIGQKKNHFVVVQSLNQGNINYNDKDFDLDLDESRYENKTAE 685
DB 233 ALVDMASERITLLANNGYPLEKHLAVYRQMLEVSLAN----QOAGSLSDATY-----KE 283
QY 686 YQEMDMKDLKLIQNKDAKIAVSIYPNTPLNGLVKKLNNIIEYNKNIFVILHVD-KNH 744
DB 284 FEMKQRLNLQSRQESSEKKAIVLAANYGYVDQVLTITKISYCHNRSIRFYLIHSDFNE 343
QY 745 LTPDIKKEILAF-----YHKQVNLNNDISYTSNRLIKTEAHLNINKLSQNL 796
DB 344 WIKOLNKRLEKFOSEIINCRTVTSQIS-CYKSDISYTVFLRYFTADF-----VQEDKALYL 398
QY 797 NCEYIIFDNHDSLFVKN-DSV-----AYMKKIDYGMN 827
DB 399 CDCULVYTKNLDLFPATLODYPLAARDVFGGRAYGQEIFNAGVLLYNNAFWKKENTQK 458
QY 828 FSALTDHWIEKIN 840
DB 459 LIDVTNEWHDKVD 471

RESULT 13
Q8XN34 PRELIMINARY; PRT; 348 AA.
AC Q8XN34;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Spore coat polysaccharide biosynthesis protein.

```

```

GN      CPE0504.
OS      Clostridium perfringens.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC      Clostridiales; Clostridiaceae; Clostridium.
OX      NCBI_TaxID=1502;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=13 / TYPE A;
RX      PubMed=11792842;
RA      Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA      Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT      "Complete genome sequence of Clostridium perfringens, an anaerobic
RT      flesh-eater";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR      EMBL; AP003187; BAB80210.1;
DR      InterPro; IPR001173; Glycos_transf_2.
DR      pfam; PF00535; Glycos_transf_2; 1.
KW      Complete proteome.
KW      SEQUENCE 348 AA; 41769 MW; AB46C19FA8E52C93 CRC64;

Query Match      5.2%; Score 265.5; DB 16; Length 348;
Best Local Similarity 23.3%; Pred. No. 2.1e-06;
Matches 109; Conservative 68; Mismatches 132; Indels 159; Gaps 18;

QY 435 VSIYIPAYNCANYIQCVCDSALNQTVWVLEVCICNDGSTDNTLEVINKLYGNPRVRIMS 494
DB 6 VSIIPAYNIESYVERCLNSLINQTYENIEIIVDGDSTDNTLVKISDMARKDKRIKYLE 65
QY 495 KPNGGIASASNAASFAGYIYGQDSDYDLEPDVAVELCLKFEFLDKTLACVYTTNRNRYN 554
DB 66 QENRGSSEARKKGEMSTGEFILFVDGDDWIRNDIEVLE-----YS-----N 109
QY 555 PDGSLIANGYNWPEFSREKLTMTAMIAHFRMFTIRAWHLTDGFNENIENAVDYDMFLKLS 614
DB 110 DDIDIVAFGYN-----EIFNEN----- 126
QY 615 EVGKFKHLNKICYNRVLHGDNTSIKKGIGQKKNHFVVVQSLNQGNINYNDKDFDLDL 674
DB 127 EIQK-----NSLLYDK-----NNEKIELRDTQFLREILTNNISIN-----IWKNF---I 167
QY 675 SRKYIFENKTAEYQEMDMKDLKLIQNKDAKIAVSIYPNTPLNGLVKKLNNIIEYNKNIF 734
DB 168 RKFFIDKNVNVFPKMSYAEADALLISLAAK-----EPNVI 203
QY 735 VIILHVDKNHLPDIKKEILAFYHKHQ-----VNILLNNDISYTSNRLIKTEAHLN 786
DB 204 VI-----KELYFYFKRENSTISISKILETKDAMFIKNILIE----- 243
QY 787 NINKLSQNLN---CEYI---IFDNHDSLFVKNDSYAYMKKYDVGMMFNSALTDHWIEKIN 840
DB 244 --NKLNIYEEFEYCYVIHNNWYFRYRIYFGNNQYS-KKLFEIWNFNFI-----SIR 293
QY 841 AHPFPFKLIKLT---YFNDNDRSMNVKVGASQGMFKYALPHELLTIK 885
DB 294 NNKYFKKDEKNLDCYWRFFVFLNAIN-----NYLKGILALIIR 331

RESULT 14
Q8YSL7 PRELIMINARY; PRT; 323 AA.
AC Q8YSL7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Alr3067.
GN Anabaena sp. (strain PCC 7120).
OS Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

```


THIS PAGE BLANK (USPTO)